



AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>



Thèse présentée pour l'obtention du titre de
Docteur de l'Université de Lorraine
Spécialité «écotoxicologie, biodiversité, écosystèmes»

Par **Grégory HOFF**

Réparation des cassures double-brin et variabilité chromosomique chez *Streptomyces*

Double-strand break repair and chromosomal variability in *Streptomyces*

Soutenance publique prévue le 13 décembre 2016

Composition du jury :

Rapporteurs

M. Claude Bruand, Directeur de recherche, INRA

M. Jean-Luc Pernodet, Directeur de recherche, CNRS

Examineurs

Mme Bénédicte Michel, Directrice de recherche, CNRS

M. Pierre Leblond, Professeur, Université de Lorraine (Directeur de thèse)

Mme Annabelle Thibessard, Maître de conférences, Université de Lorraine (Co-directrice de thèse)

Invités

M. François Lecointe, Chargé de recherche, INRA

Mme Claire Bertrand, Maître de conférences, Université de Lorraine (Co-encadrante de thèse)

UMR1128 DynAMic,
INRA-Université de Lorraine, Faculté des Sciences et Technologies,
Boulevard des Aiguillettes, BP 70239, 54506 Vandœuvre-lès-Nancy cedex.

REMERCIEMENTS

Dans un premier temps, je tiens à remercier l'ensemble des membres du jury d'avoir accepté de juger mon travail.

L'aboutissement de cette thèse n'aurait pas été possible sans l'aide et le soutien des gens qui m'ont entouré durant ces trois années. Pour cela j'aimerais sincèrement remercier mon directeur de thèse, Pierre Leblond, ainsi que mes co-directrices Annabelle Thibessard et Claire Bertrand, qui m'ont accordé leur confiance et m'ont encadré avec brio. Merci, pour votre disponibilité, vos conseils, mais aussi pour ces moments de rire qui font que l'on se sente bien dans son environnement de travail. Tout cela m'a permis d'aller de l'avant et de progresser, aussi bien dans mon travail que dans la vie de tous les jours.

Un grand Merci à Emilie, pour m'avoir initié aux joies de manipuler *Streptomyces*, mais également pour ta précieuse aide technique, ce qui m'a permis d'avancer plus rapidement dans mon travail. Merci également à Laurence, maillon essentiel de l'équipe technique et toujours disponible quand on a besoin d'elle.

Un Merci tout particulier à Charles, Nari et Maxime, collègues et amis « doctorants » qui naviguent dans la même direction que moi. Toutes ces discussions et ces moments passés ensemble me manqueront.

Merci à Aurélien, Perrine et Thomas, que j'ai eu la chance d'encadrer durant leur stage.

Merci également à mes amis ainsi qu'à ma famille. Je tenais tout particulièrement à remercier mes parents, soutien moral et financier indispensable pour ces 8 années d'études. Merci de m'avoir laissé aller là où je le souhaitais et de m'avoir accordé votre confiance, je vous en serai éternellement reconnaissant.

Enfin, un immense Merci à Anais, pour son soutien sans faille et sa compréhension. Sans toi, cela n'aurait pas été pareil.

LISTE DES ABREVIATIONS

ADN : Acide **D**ésoxy**R**ibonucléique

A-EJ : *Alternative-End Joining* ou mécanisme alternatif de jonction d'extrémités

ARN : Acide **R**iboNucléique

ATP : Adenosine **T**riphosphate

attB : site d'attachement bactérien

attP : site d'attachement plasmidique

AUD : *Amplifiable Unit of DNA* ou unité amplifiable d'ADN

BER : *Base Excision Repair* ou réparation par excision de bases

BIR : *Break Induced Replication* ou replication induite par les cassures

CO : Crossing-Over

CRISPR : *Clustered Regularly Interspaced Short Palindromic Repeats* ou courtes répétitions palindromiques groupées et régulièrement espacées

DAPI : 4',6-diamidino-2-phénylindole

dNTP : désoxyribonucléotide **t**ri-**p**hosphate

DSB : *Double-Strand Break* ou cassure double-brin

DSBR : *Double Strand Break Repair*

GOC : *Gene Order Conservation* ou ordre de conservation des gènes

Hfr : Haute fréquence de recombinaison

ICE : *Integrative and Conjugative Element* ou élément conjugatif et intégratif

JH : Jonction de Holliday

Kan^R : résistant à la kanamycine

Kan^S : sensible à la kanamycine

kb : kilobase

kGy : kilo Gray

Lig-Dom : domaine ligase

Mb : mégabase

MEB : microscope électronique à balayage

MET : microscope électronique à transmission

NAD : Nicotinamide **A**denine **D**inucleotide

NCBI : *National Center of Biotechnology* ou centre national de biotechnologie

NER : *Nucleotide Excision Repair* ou réparation par excision de nucléotides

NHEJ : *Non-Homologous End-Joining* ou jonction d'extrémités non-homologues

Nuc-Dom : domaine nucléase

oriT : origine de transfert

pb : paire de base

PCR : *Polymerase Chain Reaction*

PKS : polykétide synthétase

Pol-Dom : domaine polymérase

RH : Recombinaison **H**omologue

RI : Recombinaison **I**llégitime

rNTP: ribonucléotide tri-phosphate

ROS : *Reactive Oxygen Species* ou dérivés réactifs de l'oxygène

SDSA: *Synthesis Dependant Strand Annealing*

SSA : *Single-Strand Annealing*

SSB : *Single Strand Break*

T4SS : *Type IV Secretion System* ou système de sécrétion de type IV

TAP : *Telomere Associated Proteins* ou protéines associées aux télomères

THG : Transfert **H**orizontal de **G**ène

TIR : *Terminal Inverted Repeats* ou répétitions terminales inversées

TP : *Terminal Proteins* ou protéines terminales

UV : UltraViolet

SOMMAIRE

PREAMBULE.....	7
INTRODUCTION.....	9
I. <i>Streptomyces</i> : un modèle d'étude de la diversité génétique.....	11
1. Biologie et cycle de développement.....	11
2. Une linéarité chromosomique remarquable	12
3. Une forte variabilité inter-spécifique.....	12
4. Une distribution des gènes compartimentée	14
5. Une instabilité génétique liée aux réarrangements chromosomiques.....	14
II. Les sources de variabilité génétique chez les procaryotes.....	18
1. Les transferts horizontaux de gènes : porte d'entrée d'ADN exogène.....	18
1.1 La conjugaison : mécanisme majeur de THG chez <i>Streptomyces</i>	18
1.2 Maintien de l'ADN au sein du génome.....	22
2. La recombinaison dans la réparation des cassures double-brin	23
2.1 Formation des cassures double-brin	23
2.2 Réparation par recombinaison homologue.....	24
2.2.1 Fonctions cellulaires de la recombinaison homologue	24
2.2.2 Les différentes voies de la recombinaison homologue.....	25
2.2.3 Impact de la recombinaison homologue sur la plasticité génomique	27
2.2.4 Acteurs de la recombinaison homologue chez les bactéries.....	29
2.3 Réparation par recombinaison illégitime	35
2.3.1 Le NHEJ eucaryote : multiplicité des acteurs	35
2.3.2 Le NHEJ chez les procaryotes.....	36
2.3.3 L'alternative end joining chez les bactéries	42
2.4 Etat de l'art des systèmes de recombinaison chez les <i>Streptomyces</i>	42
2.4.1 La recombinaison homologue.....	42
2.4.2 La recombinaison illégitime.....	42
OBJECTIFS DE LA THESE.....	45
RESULTATS.....	47
I. Etude des acteurs tardifs de la recombinaison homologue.....	49
II. Caractérisation des gènes NHEJ-like chez les <i>Streptomyces</i>.....	63
III. Conséquences de la réparation des cassures double-brin sur le génome de	
<i>S. ambofaciens</i>.....	111
1. Construction d'un système engendrant des DSB.....	111
2. La réparation mutagène de DSB dans la partie conservée du chromosome est dépendante du NHEJ	113
2.1 L'analyse des clones Kan ^r révèle des cicatrices typiques d'une réparation NHEJ.....	115

2.2	Le NHEJ permet l'insertion de matériel génétique ectopique au site de cassure.....	119
2.3	<i>kuA</i> serait un acteur majeur mais non exclusif du NHEJ.....	120
2.4	L'analyse des clones Kan ^R confirme les résultats observés pour les clones Kan ^S	120
3.	La réparation des DSB dans les bras chromosomiques induit de l'instabilité génétique et génomique.....	121
3.1	Analyses des phénotypes coloniaux.....	121
3.2	Observation/analyse de l'intégrité des génomes par PCR.....	122
3.3	Analyse des événements de réparation.....	126
3.3.1	Réparation de DSB par remplacement de bras chromosomique.....	126
3.3.2	Circularisation du génome.....	128
3.3.3	L'amplification d'ADN dans le cluster de gènes de la stambomycine comme conséquence de la réparation des DSB.....	131
DISCUSSION GENERALE		135
I.	Etude des acteurs tardifs de la recombinaison homologe chez <i>S. ambofaciens</i>	137
1.	Rôle de RuvABC et RecG dans la migration et la résolution des jonctions de Holliday.....	137
2.	Possible redondance fonctionnelle des protéines intervenant dans la migration et la résolution des jonctions de Holliday.....	140
II.	Le NHEJ chez <i>Streptomyces</i>	142
1.	Importance du NHEJ dans le cycle cellulaire de la bactérie.....	142
2.	Caractérisation des gènes NHEJ- <i>like</i> chez les <i>Streptomyces</i>	144
2.1	Implication dans le NHEJ.....	144
2.2	Implication des gènes NHEJ- <i>like</i> dans d'autres mécanismes de réparation de l'ADN.....	149
III.	Prise en charge des cassures double-brin le long du chromosome	152
1.	Conséquences de la réparation de DSB sur le chromosome.....	152
1.1	Des modifications mineures liées au NHEJ dans la région core.....	152
1.2	Des réarrangements majeurs dans les bras du chromosome.....	153
1.2.1	La recombinaison homologe joue un rôle prépondérant dans les réarrangements génomiques observés.....	153
1.2.2	La recombinaison illégitime a un impact majeur sur le génome.....	158
1.3	Prise en charge des DSB chez d'autres microorganismes.....	158
2.	Existence d'un mécanisme de recombinaison illégitime alternatif.....	159
2.1	Caractéristiques du mécanisme.....	159
2.2	Acteurs possibles.....	160
3.	Régulation de l'intervention RH/NHEJ.....	162
PERSPECTIVES		167
Bibliographie.....		169
ANNEXES		185

PREAMBULE

Au cours des deux dernières décennies, l'amélioration des techniques de séquençage a permis d'obtenir pas loin de 75000 génomes dont près de 95% sont d'origine procaryote (<http://www.ncbi.nlm.nih.gov/genome/browse/>). Ces derniers comprennent des gènes qui sont partagés par l'ensemble de la population ou espèce (souvent essentiels à la croissance végétative) et qui définissent le génome 'core'. Contrairement aux eucaryotes où la structure de l'information génétique est commune entre individus d'une même espèce (aux mutations et réarrangements près), un organisme bactérien dispose en plus de cette information « core » d'une collection de gènes qui est présente uniquement dans une fraction des membres d'un clade et qui constitue le génome dit flexible ou accessoire. Ce dernier confère à un individu ou à la population entière des capacités d'adaptation aux changements environnementaux biotiques et abiotiques. Au sein d'une même espèce, la somme de tous les gènes accessoires excède le plus souvent la taille du génome d'un individu, soulignant l'extraordinaire répertoire génique dont disposent les bactéries et archées. Cette part de gènes variables résulte d'événements de gains et de pertes d'information génétique au sein d'un phylum. L'acquisition d'information génétique en provenance de congénères non ascendants est un phénomène maintenant bien connu et appelé transfert horizontal. Cela s'oppose à la transmission verticale de l'information génétique, c'est-à-dire des ascendants vers les descendants, mode d'hérédité quasi exclusif chez les organismes à reproduction sexuée.

Les phénomènes d'acquisition et de perte d'information génétique constituent des facteurs de l'évolution de la structure du chromosome, dont l'architecture est par ailleurs fortement contrainte par des mécanismes cellulaires majeurs comme la réplication, la ségrégation chromosomique ou la réparation de l'ADN. En tenant compte de ce constat, la question suivante peut alors se poser : quels sont les mécanismes qui influent sur la plasticité d'un génome bactérien ?

Les bactéries du genre *Streptomyces*, de par la structure atypique de leur génome, constituent à cet égard un modèle chromosomique très intéressant pour l'étude des mécanismes de la plasticité du génome. En effet, l'information accessoire peut représenter jusqu'à 75% du génome. Elle est majoritairement localisée au niveau des extrémités d'un grand chromosome linéaire tandis que l'information conservée est localisée au centre du chromosome. La forte proportion de gènes variables ainsi que l'organisation remarquable du chromosome soulèvent de nombreuses questions sur le rôle de ces gènes au sein de la

bactérie, de la population ou encore de l'écosystème, ainsi que sur les mécanismes qui entretiennent cette variabilité.

Au travers de ce travail de thèse, les résultats acquis ont permis d'identifier des acteurs importants de la recombinaison de l'ADN chez les *Streptomyces*, de démontrer l'existence d'un mécanisme de réparation mutagène et de montrer que la réparation d'ADN est un moteur de la diversité génomique. Ils ouvrent ainsi des perspectives sur la compréhension de l'organisation du chromosome des *Streptomyces* et sur l'impressionnante diversité dont font preuve les bactéries de ce genre.

CHAPITRE 1 :

-

INTRODUCTION

I. Streptomyces : un modèle d'étude de la diversité génétique

1. Biologie et cycle de développement

Les *Streptomyces* sont des bactéries Gram+ dont l'écosystème principal est le sol et qui appartiennent au groupe des actinomycètes. Ce genre bactérien comporte plusieurs centaines d'espèces qui sont pour la plupart saprophytes, à l'exception de *Streptomyces scabies* qui est un pathogène de la pomme de terre et *S. somaliensis* qui est un pathogène opportuniste de l'homme. Les *Streptomyces* synthétisent un large panel de métabolites secondaires dont de nombreux antibiotiques. Contrairement à la majorité des bactéries qui se multiplie par fission binaire, les *Streptomyces* présentent un cycle de développement complexe (Chater, 1993 ; **Figure I1**). La germination d'une spore uni-génomique donne naissance à un filament pluri-génomique appelé hyphe. Celui-ci se développe par croissance apicale pour ensuite se ramifier et former un réseau que l'on nomme « mycélium végétatif ». Lorsque le milieu devient limitant en nutriments, l'activation de gènes spécifiques entraîne une différenciation morphologique et permet ainsi la formation d'un mycélium aérien constitué d'hyphes non ramifiés. Des chaînes de spores se forment alors après septation des hyphes aériens. Lorsque ces spores arrivent à maturité, elles se détachent et leur dissémination permet d'initier un nouveau cycle (Chater, 1993).

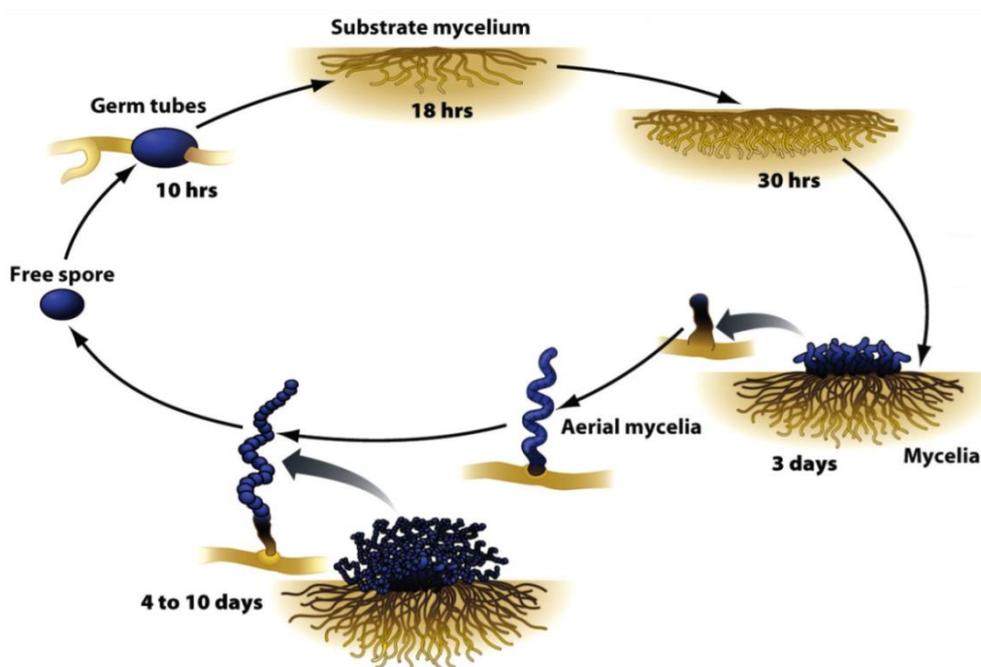


Figure I1 : Cycle de développement sur milieu solide des bactéries du genre *Streptomyces*.

Source : McGill University

2. Une linéarité chromosomique remarquable

En plus de leur cycle de vie atypique, les *Streptomyces* possèdent une organisation chromosomique remarquable. Contrairement à la majorité des bactéries, les *Streptomyces* disposent d'un chromosome linéaire dont le taux de G et C est supérieur à 70%. A ce jour, seul le génome de *Streptomyces violaceusniger* Tu4113 fait exception par sa circularité. Cette linéarité pourrait résulter de la recombinaison entre un chromosome circulaire ancestral et un plasmide linéaire (Volf et Altenbuchner, 2000). L'origine de réplication est située au centre de la molécule (*oriC*) tandis que les extrémités sont constituées de séquences terminales répétées (TIR pour *Terminal Inverted Repeats*) dont la taille et la composition varient en fonction des espèces. Ces dernières sont liées de manière covalente à des protéines terminales appelées TP (pour *Terminal Proteins*), elles-mêmes associées aux protéines TAP (pour *Telomere-Associated Proteins*) (Bao et Cohen, 2001 ; Chen *et al.* 2002). Ces protéines interagissent entre elles et forment un complexe capable de donner une structure au chromosome et ainsi maintenir sa stabilité. (Bao et Cohen, 2003).

3. Une forte variabilité inter-spécifique

Au moment de rédiger ce manuscrit, 42 génomes complets de *Streptomyces* étaient disponibles sur la base de données NCBI (National Center of Biotechnology : <http://www.ncbi.nlm.nih.gov/>). Une des particularités majeures des *Streptomyces* est la forte variabilité qui existe entre deux espèces appartenant à ce genre. La taille du chromosome peut varier du simple au double, allant de 5,96 Mb pour *Streptomyces xiamenensis* à 11,94 Mb pour *Streptomyces bingchenggensis*. Une étude de génomique comparative effectuée sur 17 génomes de différentes espèces de *Streptomyces* a révélé un pan-génome de 34 592 clusters (Kim *et al.*, 2015). Le pan-génome se définit comme l'ensemble des gènes présents au sein du phylum étudié. Il est constitué à la fois du génome « core », c'est-à-dire la partie commune à l'ensemble des espèces étudiées et du génome « flexible » qui correspond aux gènes présents uniquement dans quelques voire une seule espèce (Cordero et Polz, 2014). La taille du pan-génome de *Streptomyces* s'accroît avec le nombre d'espèces séquencées (**Figure I2-A**), ce qui permet de considérer que ce genre bactérien dispose d'un pan-génome ouvert où les gènes variables sont majoritaires. En effet, le génome « core » est constitué de 2018 clusters d'orthologues (**Figure I2-B**), ce qui représente seulement 24 à 38% de la taille des génomes étudiés. Ces valeurs contrastent avec le nombre de familles de gènes qui sont présentes dans au moins deux génomes (11743) ou qui sont spécifiques d'une seule espèce étudiée (20 831) (Kim *et al.*, 2015). La

variabilité génétique des *Streptomyces* peut également être appréciée lorsque l'on se place à l'échelle intra-spécifique. Chez *Streptomyces ambofaciens*, la comparaison entre les souches ATCC 23877 et DSM 40697 a révélé que la taille du chromosome diffère de 200 kb entre les deux souches. De plus, certains éléments génétiques mobiles qui sont présents chez ATCC 23877 ne le sont pas chez la DSM 40697 et réciproquement (Thibessard et Leblond, 2016).

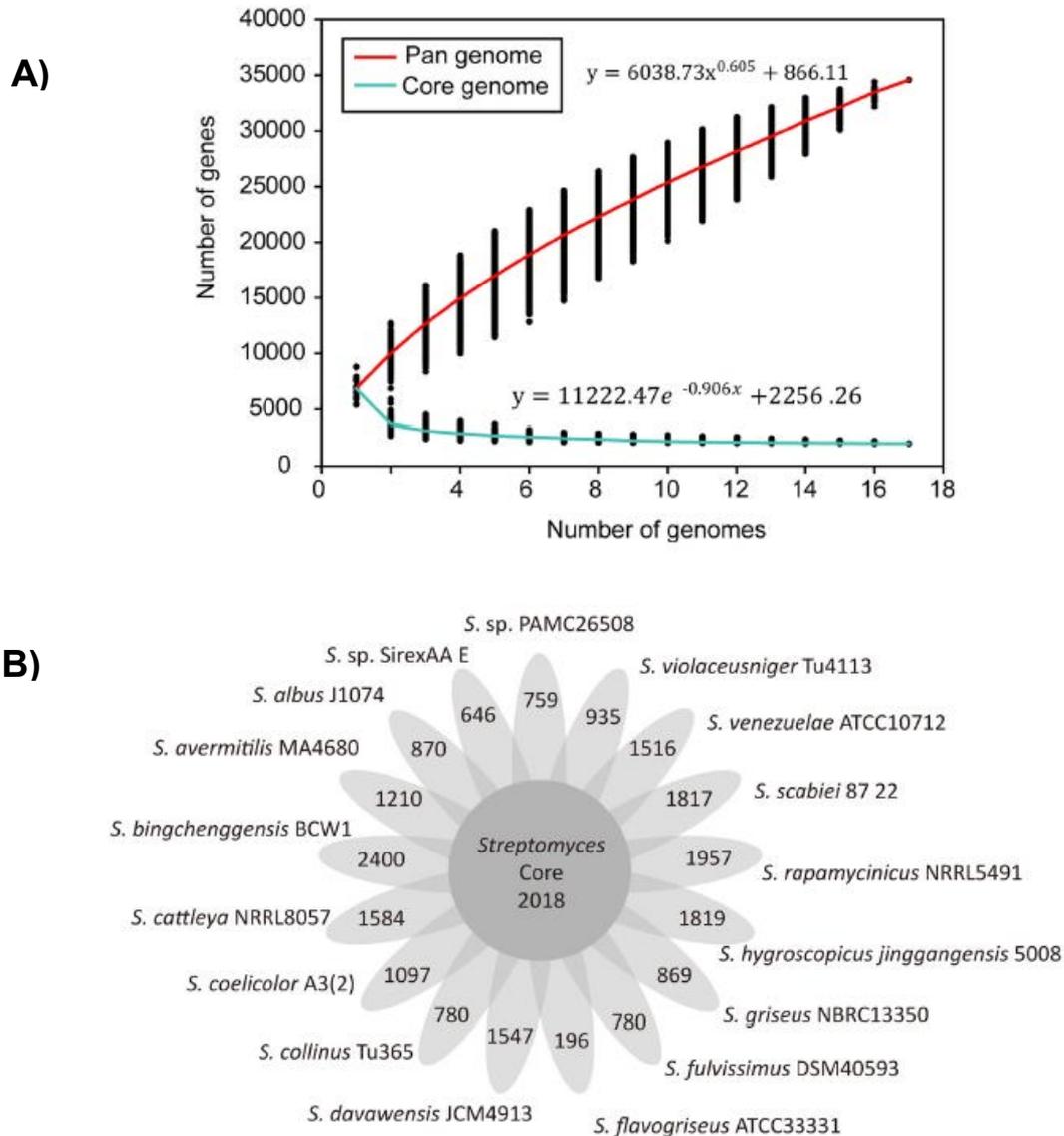


Figure I2 : Analyse du pan-génomme de 17 espèces de *Streptomyces*. **A)** Evolution du nombre de gènes constituant le pan-génomme et le core génome en fonction du nombre d'espèces étudiées. **B)** Diagramme de Venn montrant le nombre de familles de gènes spécifiques de chaque espèce pour chaque génome testé. Le nombre de gènes du core génome est représenté dans la partie centrale du diagramme (Kim *et al.*, 2015).

4. Une distribution des gènes compartimentée

Le séquençage des génomes de *Streptomyces coelicolor* et *Streptomyces avermitilis* a révélé une disposition génétique particulière : l'ensemble des gènes supposés essentiels sont situés dans la partie centrale du chromosome et forment la région « core » tandis que les gènes variables dits de contingence se trouvent majoritairement dans les bras (Bentley *et al.*, 2002 ; Ikeda *et al.*, 2003). Afin de mesurer le niveau de conservation de l'ordre des gènes le long du chromosome, une analyse comparative du génome de *S. ambofaciens* avec ceux de *S. coelicolor* et *S. avermitilis* a été réalisée en utilisant le GOC (*Gene order conservation*) (Choulet *et al.*, 2006). Le calcul de cet indice est basé sur la comparaison de génomes deux à deux et permet de déterminer le degré de synténie, c'est-à-dire la conservation d'une organisation ancestrale. En se déplaçant le long du chromosome, il s'agit de dénombrer dans une fenêtre donnée le nombre d'orthologues contigus dans les génomes comparés. Les résultats obtenus montrent une diminution progressive de la synténie à mesure que l'on s'éloigne de la région centrale du chromosome, provoquée par de nombreux événements d'insertion/délétion (indels), révélant des transferts horizontaux ainsi que des réarrangements génomiques (Choulet *et al.*, 2006).

5. Une instabilité génétique liée aux réarrangements chromosomiques

L'instabilité génétique est un phénomène largement répandu dans l'ensemble du monde vivant, mais il est présent dans des proportions particulièrement remarquables chez les *Streptomyces*. Chez *S. ambofaciens*, il a été observé que la souche sauvage engendre environ 1% de mutants affectés pour la pigmentation. Ces colonies dépigmentées sont hautement instables et sont sujettes à un phénomène nommé « hypervariabilité » (Leblond *et al.*, 1989). En effet, seulement 13% des colonies dépigmentées ont une descendance dont le phénotype est homogène. Cette instabilité qui a également été détectée chez d'autres espèces de *Streptomyces* (Birch *et al.*, 1989) est directement liée à la formation d'importants réarrangements génomiques (Birch *et al.* 1991 ; Leblond *et al.* 1990). Ces événements peuvent être de nature (circularisation, fusion, amplification...) et d'ampleur variable et affectent pour la majorité d'entre eux les bras chromosomiques.

L'analyse d'évènements de réarrangements génomiques révèle que la circularisation du chromosome intervient spontanément chez plusieurs espèces de *Streptomyces* (Fischer *et al.*, 1997 ; Kameoka *et al.*, 1999 ; Inoue *et al.*, 2003). Ces mutants se caractérisent par de grandes délétions sur les deux bras englobant les TIR et pouvant atteindre jusqu'à 2,1 Mb

soit environ un quart du génome. L'évènement de circularisation résulte d'une fusion nucléotidique illégitime entre les deux extrémités du chromosome.

L'étude de deux lignées de mutants spontanés de *S. ambofaciens* a mis en évidence une taille de TIR élevée, c'est-à-dire de 480 et 850 kb (Fischer *et al.*, 1998) contre 210 kb pour la souche sauvage (Leblond *et al.*, 1996). Ce polymorphisme des TIR résulte de la délétion d'un des deux bras entraînant son remplacement par le second. Cet échange a pu se produire grâce à un évènement de recombinaison homologue entre deux gènes codant des facteurs sigma (HasR et HasL) situés sur chaque bras du chromosome et partageant 99% d'identité nucléotidique (Fischer *et al.* 1998 ; **Figure I3**). De tels évènements ont également été observés chez *Streptomyces griseus* entre deux gènes codant des lipoprotéines (Uchida *et al.*, 2003) ainsi que chez *Streptomyces coelicolor* A3(2) entre deux transposons (Widenbrant *et al.*, 2007).

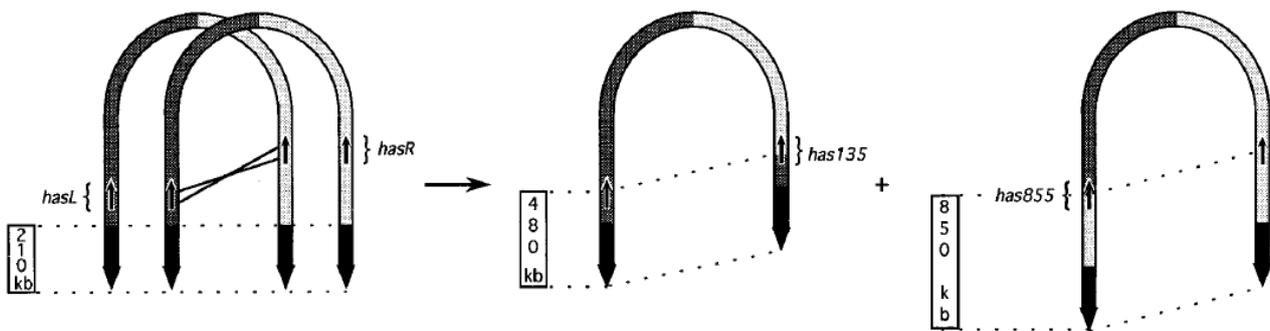


Figure I3: Représentation schématique de la recombinaison ectopique entre les loci *hasL* et *hasR* de deux copies sauvages du chromosome de *S. ambofaciens*, engendrant une première version du chromosome avec des TIR de 480 kb et une seconde version avec des TIR de 850 kb (Fischer *et al.*, 1998).

Certains mutants spontanés obtenus à partir d'une souche sauvage de *S. ambofaciens* disposent d'un chromosome dont la taille est d'environ 13 Mb contre 8,1 Mb pour la souche d'origine (Wenner *et al.*, 2003). L'analyse de l'un de ces mutants montre qu'il résulte de la fusion par recombinaison illégitime en orientation inversée de deux chromosomes délétés d'une partie du même bras (**Figure I4**). La souche obtenue est instable et produit à haute fréquence des descendants dont le chromosome a été circularisé ou dont l'un des bras a été remplacé. L'instabilité de la souche pourrait être liée à la présence d'une duplication de la région où se trouvent *parB* (qui intervient dans la ségrégation chromosomique) et *oriC* : la présence de deux origines de réplication génèrerait une forte instabilité génomique au niveau du point de rencontre des deux fourches lors de phases intenses de réplication tandis

que la duplication de la zone de ségrégation provoquerait la traction physique du chromosome vers les pôles opposés lors de la sporulation.

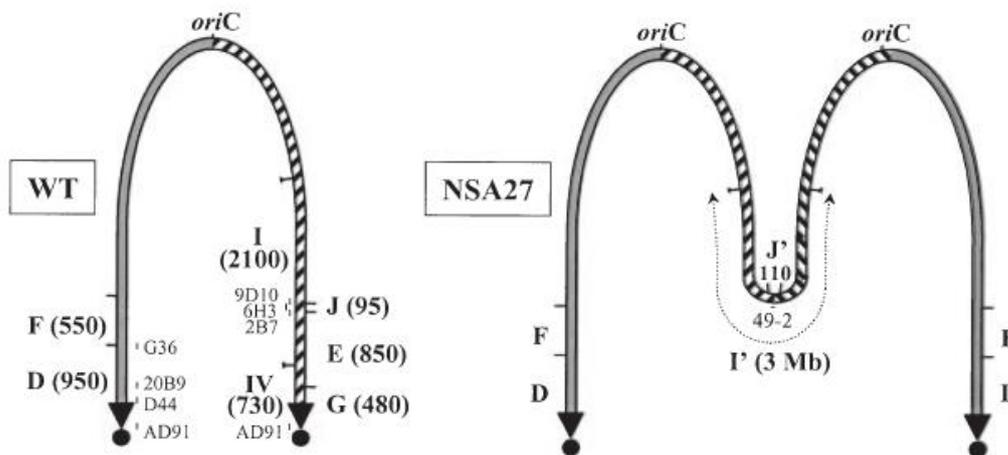


Figure 14 : Représentation schématique du chromosome du mutant NSA27 qui résulte de la fusion en orientation inversée de deux chromosomes délétés d'une partie du bras droit (Wenner *et al.*, 2003).

Chez plusieurs espèces de *Streptomyces*, des phénomènes d'amplification d'ADN chromosomique ont pu être observés. Il s'agit d'un mécanisme qui engendre la répétition en tandem d'un locus donné en plusieurs dizaines de copies. C'est un phénomène largement répandu dans le monde bactérien qui n'est donc pas spécifique des *Streptomyces*. Toutefois, il existe au moins trois caractéristiques qui semblent spécifiques aux phénomènes d'amplification observés chez les *Streptomyces* : (i) la fréquence d'apparition de cet évènement y est beaucoup plus élevée, (ii) le nombre de copies peut atteindre plusieurs centaines et (iii) le phénomène s'accompagne souvent de la formation d'une large délétion (Young et Cullum, 1987). Seuls certains *loci* particuliers, appelés unités amplifiables d'ADN (AUD), sont sujets aux amplifications (Demuyter *et al.*, 1988). Il s'agit généralement de clusters impliqués dans la production de métabolites secondaires qui comportent des gènes possédant des domaines très homologues. Pour exemple, on peut citer le cluster de synthèse de la kanamycine chez *S. kanamycetius* (Yanai *et al.*, 2006), celui de l'actinorhodine chez *S. coelicolor* (Widenbrant *et al.*, 2008) et celui des caroténoïdes chez *S. albus* (Myronovsky *et al.*, 2014).

Au sein des bactéries, les réarrangements du génome et le transfert horizontal sont des phénomènes permettant une évolution rapide. Les bactéries du genre *Streptomyces* constituent un modèle d'étude de la plasticité génomique car l'analyse de leur génome démontre que (i) les flux de gènes sont très importants au sein de l'écosystème où elles évoluent et (ii) l'organisation génomique des *Streptomyces*, c'est-à-dire la linéarité chromosomique ainsi que la localisation terminale de la variabilité suggère l'existence de mécanismes de recombinaison de l'ADN originaux le long du chromosome et/ou une tolérance aux réarrangements d'ADN dans les régions subtélomériques. Dans ce contexte, la suite de l'introduction sera focalisée d'une part sur les mécanismes de transfert de gènes chez les bactéries en se concentrant tout particulièrement sur ceux existant chez les *Streptomyces* et d'autre part sur les mécanismes de recombinaison de l'ADN qui permettent l'intégration de matériel exogène au sein d'un génome receveur et des remaniements du génome endogènes. Ce second point sera notamment axé sur la réparation des cassures double-brin (DSB), phénomène pour lesquels les mécanismes de recombinaison (recombinaison homologue, recombinaison illégitime) constituent un moyen de sauvegarde pour contrecarrer l'effet délétère de ces dommages.

II. Les sources de variabilité génétique chez les procaryotes

1. Les transferts horizontaux de gènes : porte d'entrée d'ADN exogène

Les transferts horizontaux de gènes (THG) constituent une force majeure dans la dynamique d'évolution des génomes. L'acquisition de matériel génétique conférant des fonctions nouvelles peut faciliter l'adaptation de l'organisme receveur à un nouvel écosystème. Elle permet ainsi d'améliorer la fitness de l'organisme c'est-à-dire sa valeur sélective. La propension d'acquisition de gènes par transfert horizontal fluctue en fonction des organismes et est fortement influencée par les contraintes environnementales. Par exemple, une analyse complète du génome de la bactérie autonome *Escherichia coli* MG1665 a permis d'estimer que près de 18% des gènes ont été acquis par THG depuis que l'espèce et la lignée des *Shigella* ont divergé (Lawrence et Ochman, 1998). En revanche, chez *Buchnera*, symbiote obligatoire des aphides, aucun évènement de THG n'a pu être détecté (Tamas *et al.*, 2002). A l'inverse d'*E. coli* qui se développe dans l'intestin, un écosystème très dense ou de nombreux évènements de THG ont été détectés (Liu *et al.*, 2012), la bactérie intracellulaire *Buchnera* est protégée par son habitat clos de l'opportunité d'acquérir du matériel génétique exogène.

Il existe trois mécanismes de THG : (i) la transformation naturelle définie comme étant l'acquisition par un organisme d'ADN se trouvant sous forme libre dans l'environnement, (ii) la transduction bactérienne qui consiste en un transfert de matériel génétique d'une bactérie donneuse à une bactérie receveuse par l'intermédiaire d'un bactériophage, et (iii) la conjugaison bactérienne, c'est-à-dire le transfert d'ADN par contact direct entre deux bactéries après formation d'un pore de conjugaison entre les deux cellules (**Figure I5**).

1.1 La conjugaison : mécanisme majeur de THG chez *Streptomyces*

Chez les bactéries du genre *Streptomyces*, la conjugaison bactérienne apparaît comme le mécanisme de transfert prédominant. En effet, probablement en raison d'une forte imperméabilité de la paroi, la transformation naturelle ne semble pas être possible. De plus, des évènements de transduction par les phages n'ont été décrits que très rarement, c'est-à-dire pour le phage *phiSV1* spécifique à *Streptomyces venezuelae* (Stuttard, 1979) et pour les phages DAH capables de transférer *in vitro* du matériel génétique de *S. coelicolor* vers *Streptomyces verticillum* (Burke *et al.*, 2001). En revanche, de nombreux éléments conjuguatifs réplcatifs et intégratifs ont été identifiés dans les différents génomes de

Streptomyces (pour revue : te Poele *et al.*, 2008), suggérant un rôle important de la conjugaison dans la dynamique du génome des *Streptomyces*. D'un point de vue mécanistique, la conjugaison chez les *Streptomyces* diffère cependant du modèle dépendant de l'origine de transfert (*oriT*) présent chez la grande majorité des bactéries.

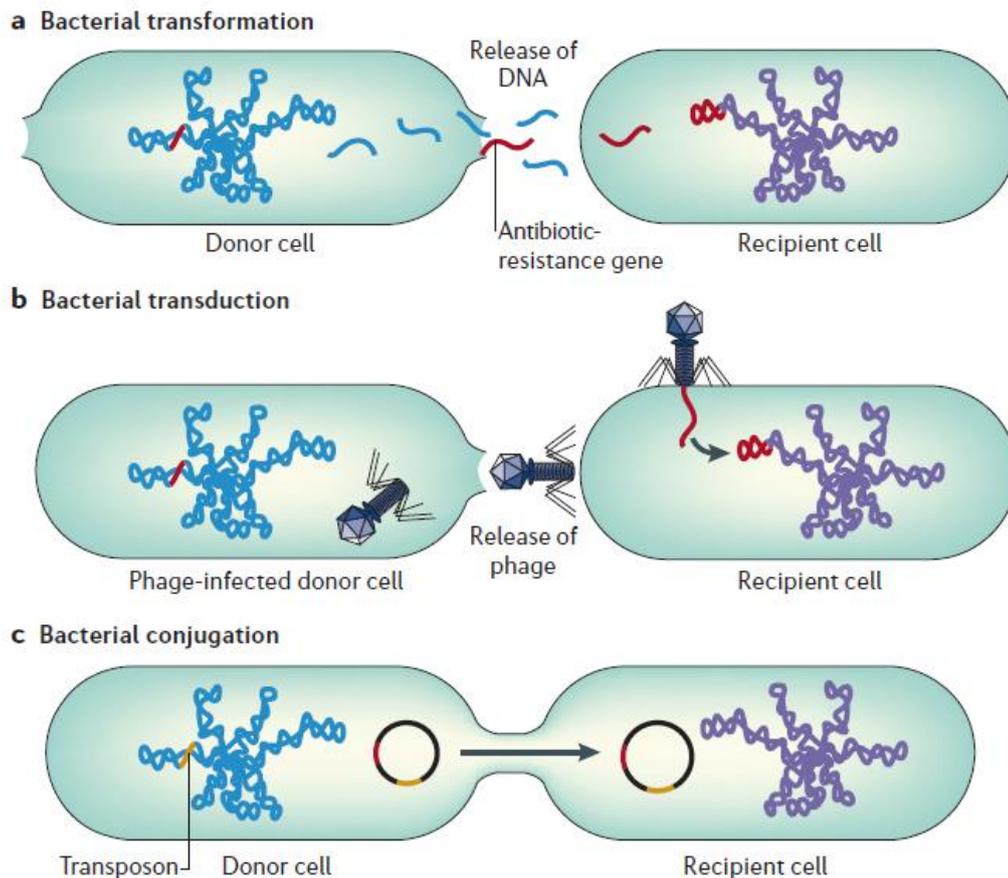


Figure I5: Représentation des différents mécanismes de transferts horizontaux de gènes : a) la transformation naturelle b) la transduction bactérienne c) la conjugaison bactérienne (Furuya et Lowy, 2006).

La conjugaison bactérienne « classique » dépendante d'*oriT*

La conjugaison bactérienne implique généralement le transfert de plasmides ou d'éléments conjugatifs et intégratifs (ICE). Comme leur nom l'indique, les ICE sont des unités d'ADN qui bien que présentes au sein du chromosome peuvent s'exciser et se transférer vers une cellule receveuse puis se réintégrer dans le chromosome (Pour revue : Wozniak et Waldor, 2010). Les caractéristiques du processus de conjugaison « classique » ont été, pour la plupart, mises en évidence chez *E. coli* mais sont valables pour la majorité des espèces bactériennes connues. La première étape consiste à établir le contact entre les deux cellules par la formation d'un pore de conjugaison assemblé grâce à un système de sécrétion de type

IV (T4SS) (pour revue : Alvarez-Martinez et Christie, 2009 ; Christie *et al.*, 2014). Une fois que le pore a été mis en place, pas moins d'une quinzaine d'enzymes procède au transfert de la molécule d'ADN sous forme simple-brin de la cellule donneuse vers la cellule receveuse. Une des molécules clé de ce transfert est la relaxase, qui induit une cassure simple-brin au niveau de l'*oriT*, se lie à l'extrémité 5' de la cassure et pilote le transfert de l'ADN à travers le pore. De la même manière dans certaines souches appelées Hfr (haute fréquence de recombinaison), la mobilisation et le transfert d'une partie de l'ADN chromosomique peut se faire en *cis* après induction par la relaxase d'une cassure simple-brin dans l'*oriT* d'un élément conjugatif intégré au sein du chromosome.

La conjugaison dépendante de TraB chez les Streptomyces

Chez les *Streptomyces*, le mécanisme est totalement différent puisqu'il nécessite uniquement la protéine TraB ainsi qu'une petite séquence non codante appelée *clt* (*cis-acting locus of transfer*) pour le transfert d'un plasmide (Thoma et Muth, 2012 ; Thoma et Muth, 2015) (**Figure I6**). La protéine TraB est une ATPase hexamérique en forme d'anneau homologue de FtsK, l'enzyme clé de la ségrégation des chromosomes lors de la division cellulaire. Elle forme un pore et transloque l'ADN à travers ce dernier après reconnaissance de la séquence *clt* du plasmide conjugatif (Thoma et Muth, 2012). La séquence *clt* diffère d'une *oriT* « classique » puisque sa reconnaissance par TraB n'entraîne pas son clivage et le transfert du plasmide s'effectue sous forme double-brin et circulaire le cas échéant (Possoz *et al.*, 2001). La nature de la séquence *clt* varie en fonction des espèces mais est toujours constituée de séquences répétées capables de former des structures secondaires (Vogelmann *et al.*, 2011). Une fois le transfert primaire du plasmide effectué, une machinerie multi-protéique composée de TraB et des protéines de dispersion Spd (Spd pour *spread* : ces protéines n'ont aucune similarité avec des protéines déjà connues), se met en place et catalyse la dispersion intra-mycélienne du plasmide (Thoma et Muth, 2015).

Contrairement au T4SS, le transfert d'ADN chromosomique n'utilise pas d'élément génétique mobile intégré dans le chromosome mais nécessite tout de même l'intervention d'un plasmide conjugatif. En fonction de la forme linéaire ou circulaire du plasmide, le mécanisme de mobilisation des séquences chromosomiques mis en place va différer. Dans le cas d'un plasmide conjugatif circulaire, TraB va reconnaître des séquences *clt-like* présentes au sein du chromosome (Thoma et Muth, 2015). Chez *S. coelicolor*, l'intégralité du génome a

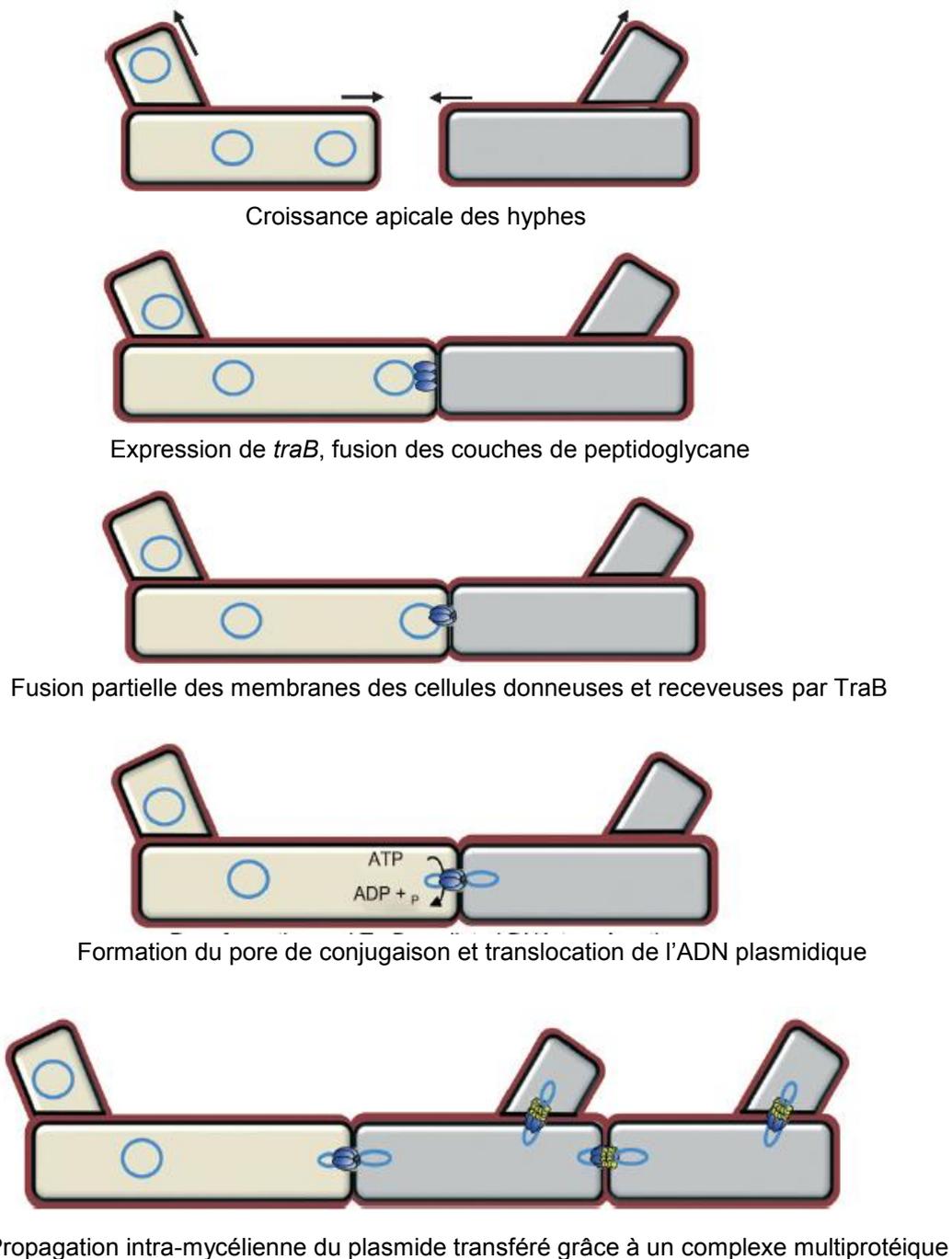


Figure 16 : Modèle de la conjugaison plasmidique chez les *Streptomyces* (Thoma et Muth 2012)

été criblée et 25 séquences *clt-like* ont pu être détectées. Des analyses *in vitro* ont permis de mettre en évidence la reconnaissance de ces séquences par TraB confortant ainsi leur rôle dans la conjugaison (Vogelmann *et al.*, 2011). Lorsque le plasmide conjugatif est linéaire, la mobilisation de l'ADN chromosomique suivrait un modèle nommé « *end-first* » (Chen, 1996). Le transfert s'initierait par les extrémités des deux réplicons linéaires qui fusionneraient pour ne former qu'une seule entité par l'intermédiaire des protéines terminales. Le transfert de la molécule hybride se ferait alors par l'intermédiaire du pore formé par TraB. La véracité de ce modèle est confortée par le fait qu'un plasmide linéaire conjugatif est capable de

mobiliser un chromosome linéaire mais pas un chromosome circularisé artificiellement tandis qu'un plasmide conjugatif circulaire est capable de mobiliser les deux (Lee *et al.*, 2011).

1.2 Maintien de l'ADN au sein du génome

Maintien à court terme

Quel que soit l'organisme concerné et le mécanisme de THG impliqué, le maintien à court terme de l'ADN exogène passe par sa réplication. Cette dernière peut se dérouler de manière autonome si la molécule dispose d'une origine de réplication, ou après intégration au sein d'un réplicon déjà présent (chromosome ou plasmide). L'intégration de l'ADN exogène peut se faire de trois manières différentes : (i) par recombinaison homologue (RH), (ii) par recombinaison illégitime ou encore (iii) par recombinaison site-spécifique. La RH peut entraîner soit un remplacement allélique, soit une co-intégration (intégration d'un réplicon dans un autre) soit l'intégration d'une information nouvelle. Elle nécessite la présence d'une homologie de séquence entre l'ADN entrant et l'ADN de la souche receveuse dont la taille varie en fonction des organismes. La recombinaison illégitime repose généralement sur le Non-homologous end joining (NHEJ), un mécanisme dépendant des protéines Ku et LigD et présent uniquement chez certaines bactéries (voir partie II.2.3). Enfin, la recombinaison site-spécifique implique l'intégration d'éléments génétiques mobiles (Phage, ICE, transposons) au sein du génome de l'hôte grâce à l'intervention d'une recombinase qui coupe au niveau de son site de reconnaissance. Pour exemple, le phage tempéré code une intégrase qui va catalyser la recombinaison entre une courte séquence d'ADN phagique (site d'attachement *attP*) et son homologue dans le génome bactérien (site d'attachement *attB*). Ce dernier peut ou non être situé dans une séquence codante (Pour revue : Groth et Calos, 2004).

Maintien à long terme

Le maintien à long terme de l'ADN transféré par THG implique une pression de sélection sur les gènes acquis. En effet, l'ADN nouvellement acquis sera conservé si : (i) il porte des gènes fonctionnels, (ii) la fonction est exprimée dans les conditions dans laquelle se trouve la bactérie, et (iii) la fonction améliore la fitness (valeur sélective) de la bactérie. Dans le cas contraire, l'ADN sera perdu par dérive génétique après accumulation de mutations et/ou délétion. Au sein des génomes bactériens, un équilibre s'établit entre acquisition et perte de matériel génétique ce qui entraîne un flux de gènes permanent.

2. La recombinaison dans la réparation des cassures double-brin

2.1 Formation des cassures double-brin

Les cassures double-brin de l'ADN ou DSB (pour **D**ouble **S**trand **B**reak) sont les dommages les plus préjudiciables que l'ADN puisse subir au sein de la cellule. Il en résulte une discontinuité dans la double hélice qui entraîne la mort cellulaire si la cassure n'est pas réparée. Les causes de formation des DSB sont très variées, et peuvent être aussi bien liées à des facteurs externes qu'à des facteurs internes à la cellule.

Au cours de son cycle, une cellule peut être exposée à différents agents clastogènes d'origine chimique ou physique. Il existe de nombreuses molécules qui engendrent des DSB de manière directe ou indirecte. Dans le cas des bactéries, ces molécules peuvent être directement produites par des organismes concurrents au sein de l'écosystème. Pour exemple, la mitomycine C est un composé synthétisé par *Streptomyces caespitosus* qui peut générer des pontages intercaténaux. La prise en charge par les mécanismes de réparation cellulaire des nucléotides affectés entraîne leur excision et induit ainsi une DSB (Mehta et Haber, 2014). La bléomycine est un composé produit par *Streptomyces verticillus* qui peut également provoquer l'apparition de DSB. C'est un composé qui est radiomimétique, c'est-à-dire qu'il va agir de manière similaire aux rayons (Mehta et Haber, 2014). Lorsque les cellules sont soumises à une exposition aux rayons ultraviolets (UV) ou aux rayons ionisants, l'ossature sucre-phosphate est attaquée par des radicaux de radiolyse entraînant ainsi la formation de cassures simple-brin (SSB pour Single-Strand Break). Lorsque la dose de rayons est élevée, la génération de deux SSB sur chaque brin complémentaire dans un même tour d'hélice devient plus fréquente ce qui augmente ainsi la probabilité d'apparition d'une DSB.

En l'absence d'un stress exogène venant perturber le cycle cellulaire, l'ADN reste tout de même très sensible aux cassures, et des DSB peuvent être provoqués par une abondance de certains composants cellulaires ou par des dysfonctionnements de la machinerie cellulaire. Les espèces chimiques oxygénées (ROS) sont des produits issus du métabolisme de l'oxygène qui sont traités par des systèmes enzymatiques variés (superoxyde dismutase, catalase...). Ces composés très actifs chimiquement peuvent se retrouver en concentration importante, et entraîner alors des dégâts pour les composants cellulaires et notamment au niveau de l'ADN. Par ailleurs, le processus de réplication est également une cause importante de génération de DSB. En effet, lors de l'avancement de la

machinerie répliquative, la rencontre avec une lésion peut entraîner un effondrement de la fourche de réplication et la génération d'une DSB (Kogoma, 1997).

Chez les eucaryotes, l'apparition de DSB peut également se faire de manière contrôlée et programmée durant certains processus physiologiques. C'est notamment le cas chez les vertébrés où le système de recombinaison V(D)J permet d'augmenter la diversité du panel d'anticorps. L'intervention de ce mécanisme se produit uniquement dans les lymphocytes et fait intervenir un mécanisme de recombinaison illégitime, le NHEJ (voir partie II.2.3) pour la réparation de DSB induites par le système lui-même (pour revue : Alt *et al.*, 1992). Outre la diversité provoquée par le « choix » des segments V, D et J au sein de multiples combinaisons possibles, la réparation par NHEJ entraîne l'ajout ou la perte d'un ou plusieurs nucléotides ce qui induit une variabilité de la séquence des anticorps. Un second phénomène, spécifique des levures, intervient lors du changement de signe conjugal. L'endonucléase HO génère une DSB au locus *mat* qui est réparée par recombinaison homologue en utilisant comme matrice un *locus* ectopique situé sur le même chromosome (Paques et Haber, 1999).

Quelles que soient l'origine et la fréquence de formation des DSB, les conséquences qu'elles entraînent rendent leur prise en charge essentielle. Au cours de l'évolution, deux mécanismes de réparation des DSB ont été sélectionnés : la recombinaison homologue (RH) et la recombinaison illégitime (RI). Ces mécanismes ainsi que leurs acteurs peuvent varier d'un organisme à l'autre. Dans tous les cas, la réparation par RH nécessite la présence d'une copie homologue intacte tandis que la réparation par RI ne nécessite pas de matrice. En fonction de l'intervention de l'un ou l'autre des mécanismes, les conséquences sur la plasticité génomique seront différentes.

2.2 Réparation par recombinaison homologue

2.2.1 Fonctions cellulaires de la recombinaison homologue

La recombinaison homologue est un mécanisme qui promeut l'appariement entre des séquences d'ADN identiques ou quasi-identiques permettant ainsi l'échange de brin d'ADN entre deux duplex de même polarité. C'est un mécanisme largement conservé dans l'ensemble des organismes vivants qui assure différentes fonctions au sein de la cellule. De manière générale, la RH est un mécanisme majeur de la réparation des DSB ainsi que des lacunes simple-brin de l'ADN. Elle est également fortement liée à la réplication et intervient notamment dans la restauration d'une fourche de réplication après arrêt de celle-ci (Pour

revue : Costes et Lambert, 2012). Chez les eucaryotes, la RH est le mécanisme moteur de la méiose (Zickler et Kleckner, 2015). Elle intervient dans le brassage génétique qui a lieu au sein du jeu chromosomique chez l'hybride. Chez les procaryotes, comme indiqué précédemment, la RH peut jouer un rôle dans l'intégration de matériel génétique exogène lors des THG.

2.2.2 Les différentes voies de la recombinaison homologue

La réparation des DSB par RH peut se faire par trois voies différentes qui coexistent au sein de la cellule : la conversion génique, c'est à dire le transfert non réciproque d'une information de séquence, le SSA (*Single Strand Annealing*) et le BIR (*Break Induced Replication*) (Rouse, 2009 ; Svendsen et Harper, 2010). En effet, après traitement du site de cassure par le biais d'une activité exonucléasique générant des extrémités 3' sortantes, la prise en charge va différer en fonction de l'origine de la cassure mais également de sa localisation. Ces différentes voies de la RH sont schématisées dans la **figure I7**.

La conversion génique (figure I7A)

Trois modèles distincts sont proposés pour la conversion génique, les étapes initiales étant communes. Après traitement du site de cassure, l'une des deux extrémités 3' sortantes nouvellement formée envahit la molécule homologue donneuse et crée une structure appelée « boucle de déplacement » (boucle D). L'extrémité envahissante agit comme une amorce pour la synthèse d'ADN et permet l'expansion de la boucle D. Parallèlement à cela, la seconde extrémité 3' sortante envahit également la matrice ce qui permet la formation d'une nouvelle boucle D qui amorce à son tour la synthèse d'ADN. Dans le modèle DSBR (*Double Strand Break Repair* : Szostak *et al.*, 1983), la progression de la synthèse d'ADN par la polymérase est possible grâce à une activité hélicase qui permet la séparation des deux brins de la molécule initiale. Cette étape que l'on nomme migration des branches résulte en la formation de deux structures à quatre brins appelées jonctions de Holliday (JH) (Holliday, 1964). La structure ainsi formée est soit résolue, soit dissoute. Dans le cas d'une résolution, une résolvasse coupe deux des quatre brins. En fonction de l'orientation des coupures par la résolvasse, le produit formé aura subi ou non un crossing-over (CO), c'est-à-dire un échange de brin entre la molécule donneuse et receveuse. Dans le modèle de dissolution de la double jonction de Holliday, il n'y a pas de résolution de la JH par une résolvasse, mais l'intervention d'hélicases et de topoisomérases déroule la structure. Le produit final ainsi obtenu est toujours généré en absence de CO. Enfin, le modèle alternatif SDSA (synthesis dependant

strand annealing) opère avant la formation des JH. Après synthèse d'ADN, la boucle D est rapidement déplacée et l'extrémité 3' néo-formée s'hybride alors avec la seconde extrémité 3' sortante qui lui est désormais complémentaire. Les lacunes sont ensuite comblées par l'action d'une polymérase. Comme pour le modèle de dissolution, il n'y a pas de formation de CO.

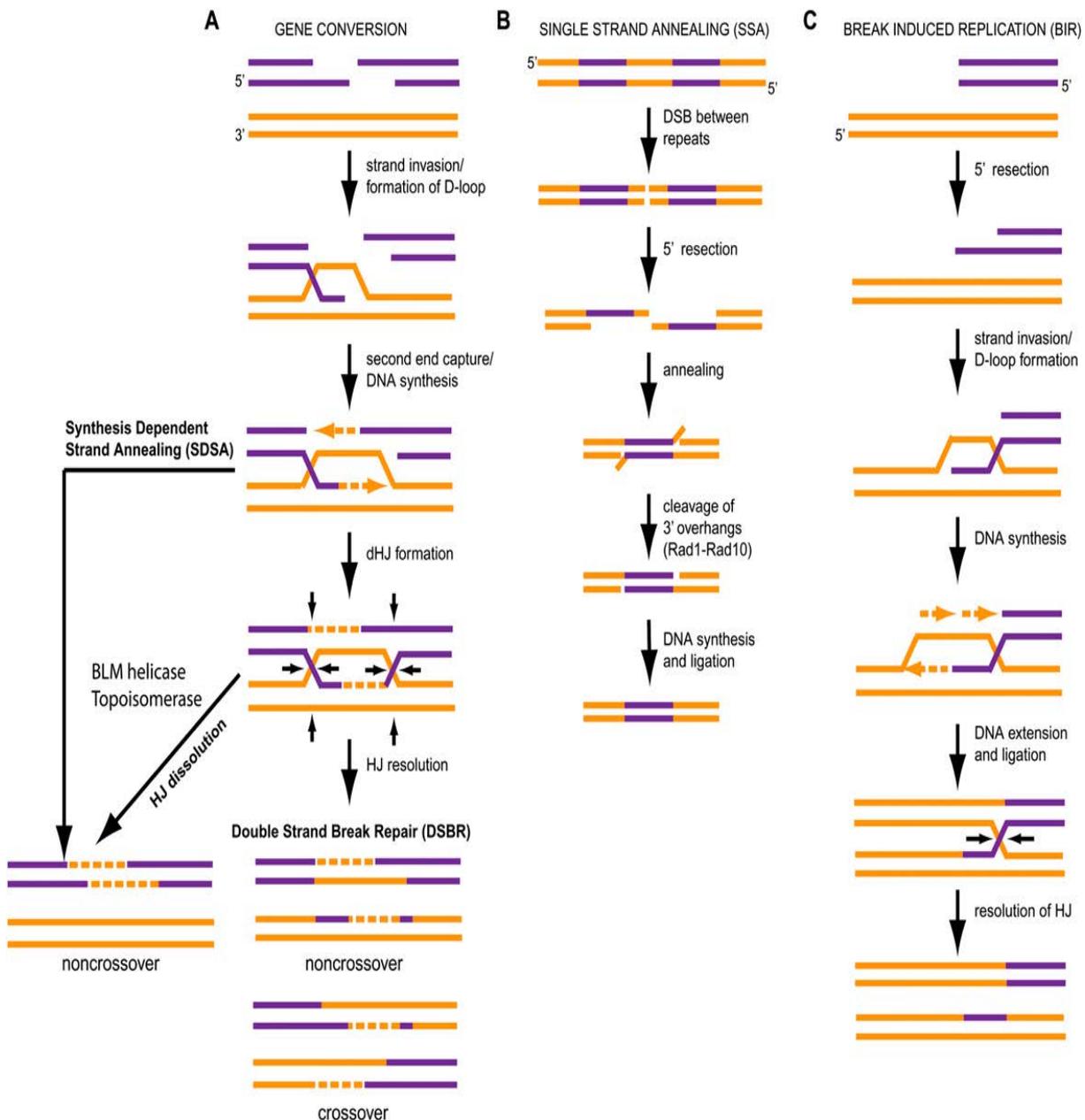


Figure 17 : Présentation des différentes voies de réparation des DSB par recombinaison homologue. **A)** La voie de conversion génique comporte trois alternatives : le DSBR (*double strand break repair*), la dissolution des doubles jonctions de Holliday et le SDSA (*synthesis dependent strand annealing*) **B)** Le SSA (*single strand annealing*) s'opère lorsque la cassure intervient entre deux séquences répétées **C)** Le BIR (*break induced replication*) s'opère lorsqu'une seule extrémité est disponible (Svendsen et Harper, 2010)

Le SSA (figure I7B)

Le SSA pour **Single Strand Annealing** est un mécanisme de recombinaison homologue qui intervient uniquement entre deux séquences répétées. Comme dans les systèmes précédents, le site de cassure est maturé afin de laisser apparaître des extrémités 3' sortantes. La séquence répétée du brin sortant située en amont de la cassure peut alors s'hybrider avec sa séquence complémentaire située sur le brin sortant en aval de la cassure. Les extrémités non hybridées sont ensuite éliminées et les lacunes sont comblées par l'action d'une polymérase. Par ce mécanisme, l'information localisée entre les deux répétitions est perdue.

Le BIR (figure I7C)

Le BIR pour **Break Induced Replication** qui est appelé RDR chez les procaryotes pour **Recombination Dependent Replication** est un mécanisme permettant la réparation d'une DSB (i) lorsqu'une extrémité du site de cassure est absente ou que la séquence nucléotidique est divergente, ou (ii) lors du redémarrage de la réplication après effondrement de la fourche de réplication. Comme pour la conversion génique, l'extrémité de la DSB est maturée afin de laisser apparaître une extrémité 3' sortante. Cette extrémité envahit alors une séquence homologue ce qui va permettre de rétablir une fourche de réplication processive entraînant l'initiation de la synthèse d'ADN. Contrairement au SDSA qui fait intervenir uniquement la synthèse du brin précoce (c'est-à-dire le brin en 5' → 3' synthétisé de manière continue), le BIR nécessite la synthèse des deux brins.

2.2.3 Impact de la recombinaison homologue sur la plasticité génomique

La RH est considérée comme un mécanisme de réparation fidèle. En effet, elle est fondamentalement conservatrice étant donné que la réparation est réalisée par copie d'une séquence homologue intacte. Pourtant, dans certains cas, l'évènement de recombinaison peut s'effectuer de manière ectopique entre deux *loci* non homologues ce qui peut entraîner des réarrangements génomiques. La fréquence de recombinaison entre deux séquences répétées dépend de plusieurs paramètres. Parmi les facteurs qui rentrent en compte, il y a notamment la taille des séquences répétées ainsi que le pourcentage d'homologie qu'elles partagent. Chez *Bacillus subtilis* la fréquence de recombinaison entre deux séquences répétées situées sur une même molécule ainsi qu'entre deux séquences répétées sur deux molécules différentes augmente de manière exponentielle avec la longueur de la séquence

identique (Michel *et al.*, 1982 ; Michel et Ehrlich, 1984). Chez la bactérie *E. coli*, une homologie de 20 paires de bases est suffisante pour obtenir une fréquence de recombinaison significative entre deux séquences répétées (Watt *et al.*, 1985). Cette efficacité augmente jusqu'à atteindre un plateau à partir de 75 nucléotides. En revanche, la présence d'une divergence nucléotidique à l'intérieur de la séquence homologue peut drastiquement diminuer la fréquence de recombinaison (Watt *et al.*, 1985). Des études similaires ont également été réalisées chez les eucaryotes. Chez la levure *S. cerevisiae* une diminution de la fréquence de recombinaison a été mise en évidence lorsque la taille des séquences homologues est inférieure à 250 pb (Jinks-Robertson *et al.*, 1985).

En fonction de la localisation et de l'orientation relatives des deux séquences subissant la recombinaison, le type de réarrangement obtenu est distinct. L'impact de l'orientation des séquences recombinantes est illustré **figure 18**.

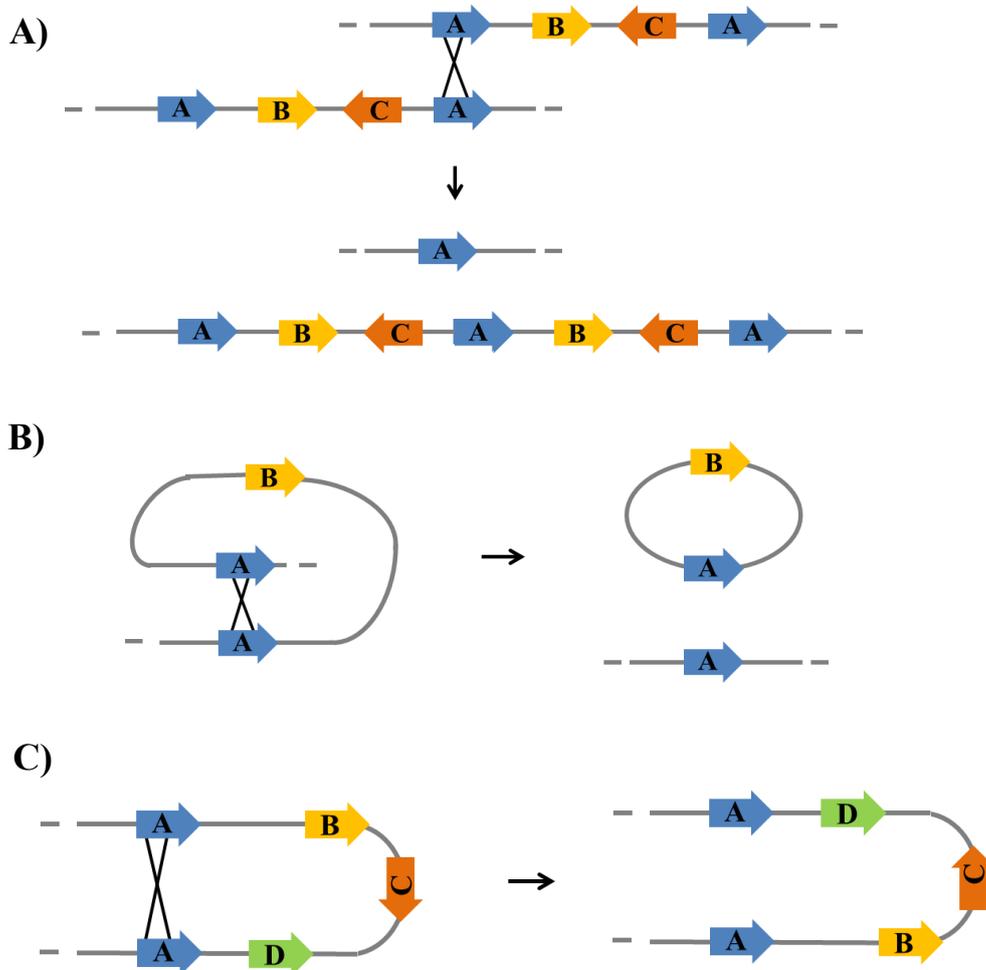


Figure 18 : Réarrangements chromosomiques résultant d'une recombinaison ectopique entre deux *loci* non homologues. **A)** Amplification/Délétion **B)** Délétion par excision **C)** Inversion

2.2.4 Acteurs de la recombinaison homologue chez les bactéries

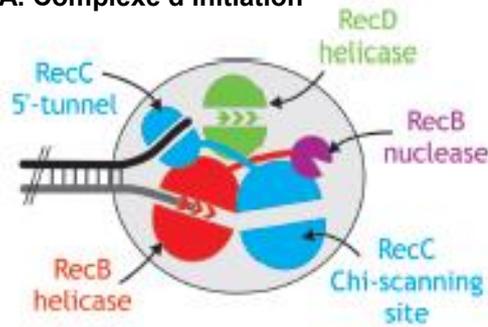
La réparation des DSB par recombinaison homologue fait intervenir de nombreux acteurs qui vont différer entre eucaryotes et procaryotes. Ce sont les bactéries *E. coli*, modèle des γ -protéobactéries, *B. subtilis*, modèle pour les firmicutes, et *Mycobacterium*, modèle pour les actinomycètes (famille à laquelle appartiennent également les *Streptomyces*) qui apportent les informations les plus complètes. Quel que soit le modèle étudié, trois étapes peuvent être distinguées : pré-synaptique, synaptique et post-synaptique.

Etape pré-synaptique

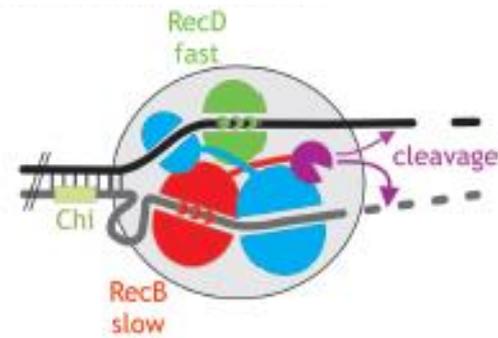
La première étape permet la maturation des extrémités de la cassure et entraîne la formation d'extrémités 3' sortantes sur lesquelles vient se fixer de façon coopérative la protéine RecA (Pour revues : Spies et Kowalczykowski, 2004 ; Dillingham et Kowalczykowski, 2008) (**Figure I9**). Chez *E. coli*, cette étape fait intervenir le complexe multienzymatique RecBCD qui déroule et dégrade l'ADN jusqu'à la reconnaissance d'une séquence spécifique de 8 nucléotides (5' GCTGGTGG 3') appelée site χ .

Le complexe est composé de trois sous-unités : (i) RecB qui dispose de domaines hélicase (rouge) et nucléase (violet) et qui peut progresser de 3' vers 5' le long de la molécule, (ii) RecC qui reconnaît le site χ (bleu) et (iii) RecD qui possède un domaine hélicase (vert) et qui progresse sur le brin complémentaire c'est-à-dire de 5' vers 3'. (Amundsen *et al.* 1990 ; Masterson *et al.*, 1992 ; Spies *et al.*, 2007) (**figure I9A**). Avant la reconnaissance du site χ , RecD joue le rôle de sous unité moteur du complexe : sa vitesse de translocation est supérieure à celle de RecB et c'est elle qui est responsable du déroulement de la double hélice (Taylor et Smith, 2003 ; Spies *et al.*, 2007) (**figure I9B**). Le domaine nucléase de RecB catalyse la dégradation des deux brins de l'ADN. Cependant il est positionné idéalement pour cliver à intervalles réguliers l'extrémité 3' sortante (gris) et moins fréquemment l'extrémité 5' sortante (**figure I9B**). Après reconnaissance du site *Chi* par RecC, un changement de conformation du complexe entraîne une diminution de la vitesse de translocation de la sous-unité RecD et un arrêt du clivage de l'extrémité 5', ce qui a pour conséquence une augmentation de la fréquence de clivage de l'extrémité 3' (**figure I9C**). La différence de vitesse entre les deux sous-unités permet la formation d'une boucle simple-brin 3' au niveau de la sous-unité RecB (**figure I9C**). Le recrutement de RecA sur l'extrémité 3' peut alors débuter.

A. Complexe d'initiation



B. Avant la reconnaissance du site Chi



C. Après la reconnaissance du site Chi

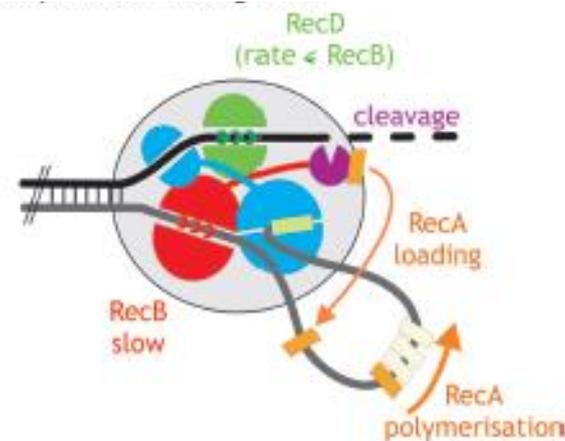


Figure 19 : Modèle illustrant la maturation des extrémités issues d'une DSB par le complexe RecBCD (Dillingham et Kowalczykowski, 2008)

Bien que le complexe RecBCD joue un rôle important chez *E. coli*, il est loin d'être ubiquiste dans le domaine bactérien (Rocha *et al.*, 2005). Chez la bactérie modèle des Gram-positives *B. subtilis*, RecBCD est remplacé par un homologue fonctionnel : AddAB (Chedin *et al.*, 1998 ; Chedin *et al.*, 2000 ; Chedin *et al.*, 2006). La sous-unité AddA est homologue à RecB, mais AddB n'est homologue ni de RecC ni de RecD (Chedin *et al.*, 2000). Contrairement à RecBCD, l'hétérodimère AddAB possède un unique domaine hélicase situé sur AddA qui pourrait à lui seul être responsable du déroulement de la double-hélice. En revanche, AddA et AddB disposent chacune d'un domaine nucléase leur permettant de cliver le brin sur lequel elles sont chargées (Yeeles et Dillingham, 2007). Le complexe reconnaît également un

site χ de 5 paires de bases (5' AGCGG 3'), différent de la séquence reconnue par *E. coli* (Chedin *et al.*, 1998).

Plus récemment, un troisième complexe hélicase-nucléase portant le nom d'AdnAB a été mis en évidence chez les mycobactéries (Sinha *et al.*, 2009). Bien que peu de choses soient connues sur cette nouvelle famille, la structure primaire d'AdnAB suggère que chaque unité du complexe soit capable de dérouler et de couper un brin de l'ADN. En effet, AdnA et AdnB possèdent chacune un domaine hélicase ainsi qu'un domaine nucléase (Sinha *et al.*, 2009). En revanche, aucun site χ n'a jusqu'à présent été identifié pour AdnAB.

Etape synaptique

L'étape synaptique est considérée comme l'étape clé de la RH (Pour revue : Cox 2007, Michel et Leach 2012). Elle est dépendante de RecA, protéine ubiquiste présente chez l'ensemble des bactéries non symbiotiques (Rocha *et al.* 2005). En se fixant sur les extrémités 3' sortantes nouvellement générées, RecA engendre la formation d'un filament nucléoprotéique hélicoïdal capable de piloter la recherche d'une séquence homologue intacte. Le nucléofilament envahit alors le duplex et s'hybride avec sa séquence complémentaire. Cela entraîne la formation d'une structure intermédiaire composée de quatre brins.

Etape post-synaptique

Le complexe RuvABC pour la résolution des jonctions de Holliday

Après invasion du brin complémentaire par RecA puis formation des boucles D, la migration de la structure entraîne la formation des jonctions de Holliday. Chez *E. coli*, les protéines RuvA, RuvB et RuvC forment un complexe tripartite (RuvABC) appelé résolvasome capable de catalyser la migration et la résolution des JH (**figure I10**). La protéine RuvA est un tétramère qui se fixe à l'ADN grâce à des domaines hélice-tige-boucle-hélice se liant au phosphate de l'ADN (Hargreaves *et al.*, 1999 ; Rafferty *et al.*, 1998). Elle permet la séparation des brins au niveau de la JH. La protéine RuvB, une hélicase hexamérique poursuit ensuite la migration (Parsons *et al.*, 1995). Enfin, l'endonucléase dimérique RuvC entraîne la résolution des JH en générant une coupure sur deux des quatre brins. Après ligation, la continuité de la double hélice est alors restaurée (Sharples *et al.*, 1999). En fonction de l'orientation des coupures (brins internes ou brins externes de la

structure), la résolution de la croix de Holliday entraîne ou non la formation d'un crossing over (CO). (Dunderdale *et al.*, 1991 ; Iwasaki *et al.*, 1991).

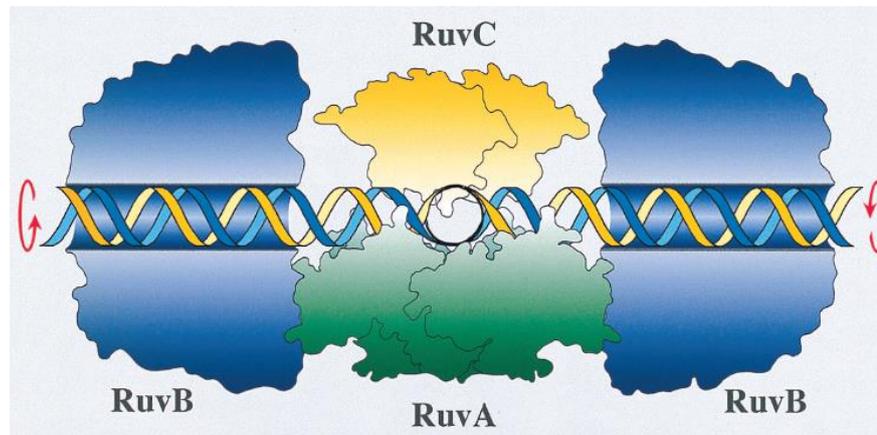


Figure I10 : Modèle représentant le résolvasome RuvABC (Sharples *et al.*, 1999).

A l'exception de rares espèces bactériennes dont quelques symbiotes obligatoires, les gènes *ruvA* et *ruvB* sont retrouvés dans une grande majorité des génomes procaryotes (Rocha *et al.*, 2005). Ces deux gènes sont toujours co-occurents et sont organisés le plus souvent sous forme d'un opéron (Rocha *et al.*, 2005). En revanche, la présence de *ruvC* dans les génomes bactériens est beaucoup moins fréquente que celle de *ruvAB*. Ce dernier peut soit être isolé ou former un opéron avec *ruvAB* (Rocha *et al.*, 2005). Chez les firmicutes incluant notamment le genre *Bacillus*, RuvC est remplacée par un analogue fonctionnel, RecU qui n'est pas homologue de RuvC (Ayora *et al.*, 2004 ; Sanchez *et al.*, 2005). La protéine RecU purifiée a une grande affinité pour les structures à trois ou quatre brins et peut cliver des JH *in vitro* en présence de MgCl₂ (Ayora *et al.*, 2004). Chez *Bacillus*, RuvAB fonctionne avec RecU de la même manière que le complexe RuvABC chez *E. coli* (Sanchez *et al.*, 2005).

RecG : une protéine pléiotrope capable de résoudre les jonctions de Holliday

Lors de la réponse aux dommages à l'ADN chez *E. coli*, l'inactivation du résolvasome RuvABC peut être compensée par la protéine RecG (Lloyd et Buckman, 1991). RecG est une hélicase capable de cibler une variété de structures complexes d'ADN (intervenant notamment dans la réplication ou le maintien des télomères) parmi lesquelles on retrouve les JH (Lloyd and Sharples, 1993a ; Lloyd and Sharples, 1993b ; Whitby *et al.*, 1993). Elle possède une activité de migration des branches similaire à celle de RuvAB et peut donc intervenir dans la réparation des DSB par RH. A l'exception des mollicutes et des chlamydiae, le gène *recG* est fortement conservé au sein des génomes bactériens (Rocha *et al.*, 2005). Bien que la protéine RecG soit capable de dérouler les JH, elle ne possède toutefois

pas d'activité nucléase nécessaire à leur résolution (Meddows *et al.*, 2004). Chez *E. coli* K-12, le gène *rusA* pourrait coder une endonucléase homodimérique capable de cliver les JH qui pourrait donc constituer une alternative à RuvC. Toutefois, le gène est situé sur un prophage et n'est pas exprimé car privé de séquence promotrice. Seule l'insertion d'une séquence incluant un promoteur (IS2 ou IS10) en amont du gène peut induire la transcription du gène (Mahdi *et al.*, 1996). Une famille d'endonucléases, comprenant notamment le gène *yqgF* chez *E. coli*, partage des propriétés structurales avec la famille de *ruvC*, et pourrait constituer une autre alternative pour la résolution des JH (Aravind *et al.*, 2000). Enfin, une dernière hypothèse suggère que le monomère RecG est capable d'agir seul, et d'éliminer la structure en menant la migration des JH de manière à leur permettre de fusionner avec une fourche de réplication (Wardrope *et al.*, 2009). Le modèle proposé par Wardrope et ses collègues est présenté dans la **figure I11**. Ainsi, lorsqu'une JH est présente dans une cellule en cours de réplication, elle peut entraver l'avancée de deux fourches de réplication indépendantes qui progressent en orientation inverse de la JH. La rencontre entre la JH et les deux fourches de réplication entraîne un échange de bras entre les deux chromosomes en cours de synthèse qui permettra la résolution de la structure après achèvement de la réplication.

Le rôle du résolvasome RuvABC au sein de la cellule paraît relativement clair tandis que celui de RecG reste à ce jour toujours vague (Lloyd et Rudolph, 2016). En effet, la protéine RecG est capable grâce à son activité hélicase de dérouler de nombreuses structures secondaires (exemple de la boucle R, une structure à trois brins composée d'ADN et d'ARN) aussi bien chez *E. coli* (Vincent *et al.*, 1996 ; Whitby and Lloyd, 1998 ; McGlynn et Lloyd, 1999) que chez d'autres bactéries telles que *Mycobacterium* (Zegeye *et al.*, 2012). Elle intervient également dans la régulation de la réplication dépendante de la recombinaison (RDR). Le redémarrage d'une fourche de réplication stoppée prématurément est ainsi possible par l'intermédiaire de PriA, une protéine qui intervient dans une forme de réplication indépendante de DnaA pouvant s'initier loin de l'origine de réplication (Kogoma *et al.*, 1997). RecG intervient également dans la ségrégation des chromosomes aussi bien chez *E. coli* que chez *Bacillus* (Sanchez *et al.* 2007 ; Rudolph *et al.* 2009a ; Rudolph *et al.* 2009b). Récemment, il a été suggéré que lors de la RH, RecG agirait préférentiellement sur la boucle D formée par le nucléofilament sur lequel est chargé RecA plutôt qu'au niveau des JH. Elle favoriserait le chargement de PriA et entrainerait ainsi une reprise efficace de la réplication (Azeroglu *et al.*, 2016). La migration et résolution des JH et le redémarrage de la réplication à partir de la boucle D seraient donc deux moyens pour stabiliser les intermédiaires de recombinaison.

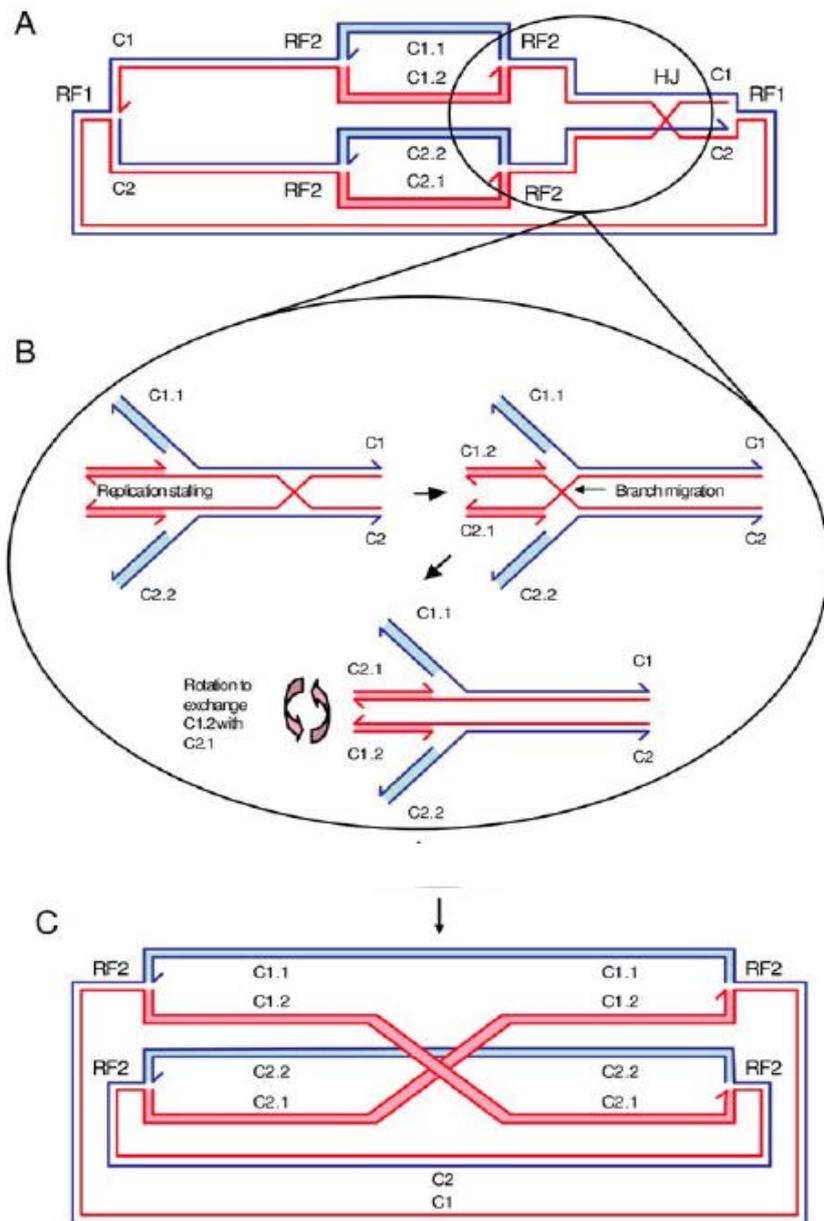


Figure I11 : Modèle représentant la maturation des jonctions de Holliday sans étape de clivage/ligation. En phase exponentielle de croissance, un nouveau cycle de réplication du chromosome d'*E. coli* peut être initié avant que le précédent arrive à terme, et donc avant l'étape de division cellulaire. **A)** Une JH s'est formée après le passage de la fourche de réplication RF1 dont résulteront les chromosomes C1 et C2. Une nouvelle paire de fourche de réplication (RF2) est initiée à partir des chromosomes C1 et C2 en cours de réplication. La machinerie réplique engendrera ainsi les chromosomes C1.1 et C1.2 à partir de C1 et les chromosomes C2.1 et C2.2 à partir de C2. **B)** Lorsque les fourches de réplique progressent en direction de la JH, cette dernière peut entraver leur progression. A ce stade, la migration des branches peut entraîner une permutation de brins entre les chromatides sœurs nouvellement formées. Les chromosomes C1.1 et C2.1 seront alors reliés à la partie non répliquée du chromosome C1 et les chromosomes C1.2 et C2.2 à la partie non répliquée du chromosome C2. **C)** La fourche RF1 a achevé son cycle de réplication tandis que la machinerie réplique s'est réassemblée au niveau des fourches RF2 et continue son cycle après avoir franchi la JH. Lorsque le cycle de réplication s'achève, il en résulte la formation de deux chromosomes monomériques (sans crossing-over, représenté en bleu) et d'un chromosome dimérique (avec crossing over, représenté en rouge) (Wardrope *et al.* 2009).

Acteurs alternatifs

Chez *E. coli*, la voie RecBCD prend en charge la réparation des DSB tandis que la voie RecFOR permet d'initier la RH au niveau de lacunes simple-brin. Cependant, lorsque l'inactivation de *recBCD* est accompagnée de mutations affectant directement ou indirectement l'exonucléase I, la voie RecFOR peut également initier la RH à partir d'une DSB (Spies et Kowalczykowski, 2004). Dans un premier temps, l'hélicase RecQ déroule la double hélice ce qui génère de l'ADN simple-brin. Les protéines SSB se fixent alors sur la molécule d'ADN et forment ainsi un ensemble SSB-ADN simple-brin. Ce complexe est la cible principale de l'exonucléase I, ce qui explique la nécessité de son inactivation pour favoriser la prise en charge des DSB par la voie RecFOR. En association avec RecQ, la protéine RecJ traite ensuite les extrémités cassées grâce à son activité exonucléase 5'→3' (Spies et Kowalczykowski, 2004). L'action concertée des protéines RecF, RecO et RecR permet alors de charger la protéine RecA sur le filament simple-brin nouvellement formé et d'initier l'étape synaptique telle qu'elle se déroule dans la voie RecBCD.

2.3 Réparation par recombinaison illégitime

Dans certaines phases du cycle cellulaire, (phase G1 chez les eucaryotes, phase stationnaire chez les bactéries) ou dans certains tissus ou cellules spécialisés comme les spores bactériennes, la réparation d'une DSB par RH est peu fréquente puisqu'un seul chromosome est présent dans chacune des cellules. Dans ces cas-là, le mécanisme principalement utilisé est le NHEJ qui joint les deux extrémités de la cassure après les avoir ou non modifiées. Le terme NHEJ a été introduit pour la première fois chez la levure (Moore et Haber, 1996). Il était alors défini comme l'ensemble des événements de réparation de DSB se produisant en absence de matrice homologue. Aujourd'hui, cette terminologie a évolué et se focalise sur le système impliquant des acteurs identifiés comme Ku70/80 et la ligase IV chez les eucaryotes et Ku et LigD chez les procaryotes.

2.3.1 Le NHEJ eucaryote : multiplicité des acteurs

Lorsqu'une DSB se produit au sein d'une cellule, celle-ci est reconnue par Ku70/Ku80, un hétérodimère toroïdal (en forme d'anneau) qui se fixe au niveau des extrémités de l'ADN. Ku70 et Ku80 sont codées par deux gènes issus de la duplication d'un gène ancestral (Doherty *et al.*, 2001). Chez les mammifères, après fixation à l'ADN, Ku70/80 recrute la sous-unité catalytique DNA-PKcs et forme un complexe stable avec cette dernière

(Lieber *et al.*, 2003). Ce complexe appelé DNA-PK s'auto-phosphoryle ce qui entraîne un changement de sa conformation et permet le recrutement de composants supplémentaires. Parmi ces composants, la protéine ARTEMIS possède plusieurs activités nucléasiques activées après phosphorylation par DNA-PK qui permettent la résection des extrémités de la cassure, c'est-à-dire une délétion nucléotidique progressive à partir des extrémités générées par la DSB. Les polymérases μ et λ peuvent synthétiser de nouveaux nucléotides en présence ou en absence de matrice. La réparation de la cassure est achevée par le complexe XLF/XRCC4/ligase IV qui lie les extrémités entre elles. L'ordre de recrutement des composants du système n'est pas figé et la diversité de compétences des différents acteurs permet d'obtenir des cicatrices de réparations variées (Lieber *et al.*, 2008).

Chez la levure *Saccharomyces cerevisiae*, le système fonctionne de manière semblable, mais certains acteurs sont différents (pour revue : Dudasova *et al.*, 2004 ; Lieber *et al.*, 2010). Après fixation de Ku70/Ku80, le recrutement du complexe Mre11/Rad50/Xrs2 permet d'effectuer un pontage entre les extrémités. L'action de la nucléase Rad27 et de la polymérase Pol4 entraîne le traitement des extrémités pour faciliter la réparation. Enfin, la ligase Lig4 accompagnée de son co-facteur Lif1, (homologues fonctionnels de la ligase IV et XRCC4 des mammifères) lie les extrémités entre elles. Le **tableau I1** récapitule le rôle des principaux acteurs du NHEJ eucaryote.

Tableau I1 : Principaux acteurs du NHEJ chez les eucaryotes

Activité	Organisme	
	<i>Saccharomyces cerevisiae</i>	Mammifères
Reconnaissance de la cassure	Ku70/Ku80	
Pontage des extrémités	Mre11/Rad50/Xrs2	DNA-PK
Nucléase	Rad27	Artemis/DNA-PKcs
Polymérase	Pol4	Pol μ /Pol λ

2.3.2 Le NHEJ chez les procaryotes

Identification des acteurs

Au début des années 2000, des gènes codant des homologues des protéines Ku eucaryotes ont été identifiés chez plusieurs espèces bactériennes (Aravind et Koonin, 2001 ; Doherty *et al.*, 2001). Quelques années plus tard, une analyse systématique des génomes

bactériens entièrement séquencés révèle une distribution des homologues *ku* sporadique au sein des clades (Rocha *et al.*, 2005). En effet, des homologues ont été identifiés chez des firmicutes, des protéobactéries ainsi que des actinobactéries. Les protéines Ku bactériennes sont composées d'approximativement 270 à 350 acides aminés (Doherty *et al.* 2001). Leur découverte a permis de disséquer *a posteriori* leurs homologues eucaryotes en trois domaines distincts : un domaine von Willebrand A (vWA) situé en N-terminale qui intervient probablement dans l'interaction protéine-protéine, une région *Ku-core* située au centre de la protéine et un domaine hélice-extension-hélice (HEH) en C-terminale qui représente un domaine de fixation à l'ADN nommé SAP (Aravind et Koonin, 2001) (**figure I12**). Seul le domaine 'core' est présent chez toutes les protéines Ku procaryotes.

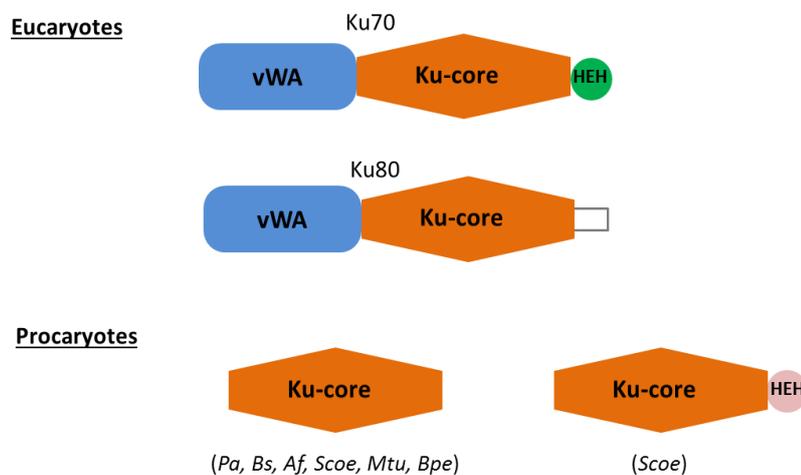


Figure I12 : Architecture des domaines des différentes protéines Ku. Les abréviations d'espèces sont les suivantes : *Pa* = *Pseudomonas aeruginosa* ; *Bs* = *Bacillus subtilis* ; *Af* = *Archeoglobus fulgidus* ; *Scoe* = *Streptomyces coelicolor* ; *Mtu* = *Mycobacterium tuberculosis* ; *Bpe* = *Bordatella pertussis*. (D'après Aravind et Koonin, 2001)

Une grande partie des génomes possédant un système NHEJ putatif ne dispose que d'un seul allèle *ku*. Cependant, chez certaines bactéries comme *Agrobacterium tumefaciens*, jusqu'à 3 homologues du gène ont été identifiés (Rocha *et al.*, 2005). Dans la majorité des cas, *ku* forme un opéron putatif avec un gène codant une ligase ATP-dépendante, ce qui laisse à penser que l'ensemble forme un système de réparation des DSB encore non défini chez les procaryotes. Chez *Bacillus subtilis* l'importance des protéines *Ku-like* YkoV et ligase-like YkoU dans la réponse aux agents génotoxiques suggère qu'elles jouent un rôle dans la réparation des DSB (Weller *et al.*, 2002).

Le modèle Mycobacterium : un système à deux acteurs

En raison de sa simplicité apparente, le système de *Mycobacterium* a très vite été considéré comme un modèle du NHEJ chez les bactéries. Chez *Mycobacterium tuberculosis*, deux acteurs, Ku et LigD, sont nécessaires et suffisants pour assurer la réparation d'une DSB (Della *et al.*, 2004). Cette dernière est une protéine multifonctionnelle composée à la fois d'un domaine polymérase (Pol-Dom), d'un domaine ligase (Lig-Dom) et d'un domaine nucléase (Nuc-Dom) (**figure I13**) (Gong *et al.*, 2005).

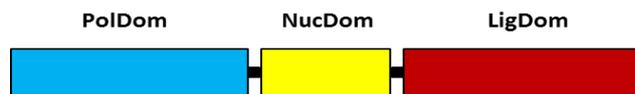


Figure I13 : Organisation des domaines de la ligase D de *Mycobacterium tuberculosis*. (D'après Pitcher *et al.*, 2007a)

Afin de mettre en évidence le rôle des différents domaines de LigD, leur activité présumée a été testée *in vitro*. Ainsi, il a été montré que Pol-Dom possède (i) une activité terminale transférase qui catalyse l'ajout de nucléotides aux extrémités d'une molécule d'ADN en absence de matrice et (ii) une activité qui permet de combler des lacunes en utilisant le brin complémentaire comme matrice (Pitcher *et al.*, 2005). Il a été constaté que pour ces deux activités enzymatiques, Pol-Dom intégrerait préférentiellement des rNTP plutôt que des dNTP, ce qui pourrait notamment s'expliquer par une plus grande abondance de rNTP que de dNTP en phase stationnaire de croissance, phase au cours de laquelle le NHEJ serait le mécanisme principal de prise en charge des DSB (Pitcher *et al.*, 2007b). Toutefois, la manière dont la cellule tolère ces ribonucléotides reste inconnue. En plus de son rôle de polymérase, Pol-Dom peut également entraîner une dissociation de l'ADN permettant de réaligner deux extrémités et d'entraîner l'appariement de nucléotides complémentaires (Pitcher *et al.*, 2007a). Enfin, l'interaction de Pol-Dom avec Ku a permis de mettre en évidence son importance dans le recrutement de LigD (Pitcher *et al.*, 2005). Lors des premières expériences effectuées sur le NHEJ chez *Mycobacterium*, il a été démontré que LigD possède une activité exonucléasique de 3' vers 5' (Della *et al.*, 2004). Toutefois, aucune activité exonucléasique n'a été détectée sur de l'ADN linéaire lorsque Nuc-Dom a été testé seul (Pitcher *et al.*, 2005). Bien que ces résultats puissent être liés au fait que Nuc-Dom seul ne puisse pas se fixer à l'ADN, son rôle dans le NHEJ reste aujourd'hui encore flou. Enfin, il a été montré que Lig-Dom peut catalyser la ligation de fragments d'ADN à extrémités compatibles (Pitcher *et al.*, 2005).

Lorsque la fonctionnalité du NHEJ a été testée, il a été démontré qu'en présence de Ku et LigD, la circularisation de plasmides linéarisés ou la réparation des DSB peut se faire de

manière mutagène (Gong *et al.*, 2004 ; Stephanou *et al.*, 2007 ; Aniukwu *et al.*, 2008 ; Gupta *et al.*, 2011). Le traitement des extrémités cassées peut entraîner des délétions uni- ou bidirectionnelles pouvant aller jusqu'à quelques centaines de nucléotides et pouvant s'accompagner d'une synthèse de nucléotides en présence ou en absence de matrice (ajout de nucléotides au niveau des extrémités ou comblement de lacunes simple-brin). Dans certains cas, la ligation de la cassure peut être facilitée par la présence de microhomologies (généralement 1 à 6 nucléotides) (Stephanou *et al.*, 2007). En fonction de la nature de la cassure, la prise en charge des extrémités sera différente. Pour des extrémités à bords francs, la réparation s'accompagne préférentiellement de synthèse de nucléotides en absence de matrice tandis que pour des extrémités 5' sortantes, la résection des extrémités est plus fréquente (Aniukwu *et al.*, 2008).

L'ensemble de ces résultats a conduit à la construction d'un modèle illustrant le NHEJ chez *Mycobacterium* détaillé dans la **figure I14**. Ainsi, la reconnaissance d'une cassure se fait par l'homodimère Ku qui recrute la protéine multifonctionnelle LigD pour le traitement des extrémités et la ligation permettant la réparation.

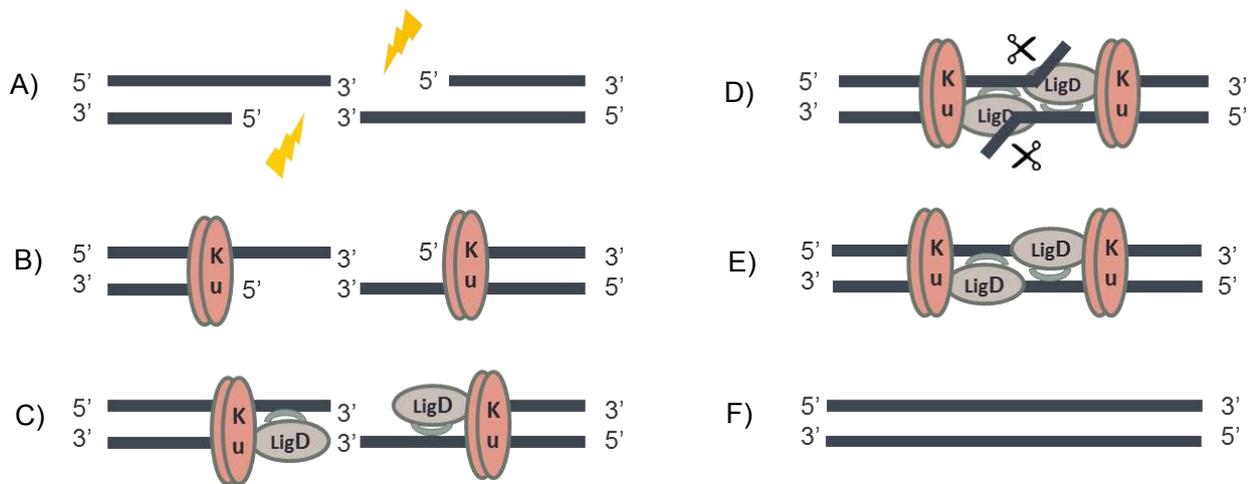


Figure 14: Représentation schématique de la réparation d'une DSB par NHEJ chez *Mycobacterium*. **A)** Formation d'une DSB **B)** L'homodimère Ku se lie aux extrémités de l'ADN **C)** Ku recrute LigD par l'intermédiaire de Pol-Dom **D)** Traitement des extrémités par perte processive de nucléotides résultant d'une activité exonucléasique et/ou par l'ajout de nucléotides par l'activité de synthèse de Pol-Dom **E)** Ligation des extrémités par Lig-Dom **F)** Réparation de la cassure. (D'après Brissett *et al.*, 2009 ; Pitcher *et al.*, 2007a).

Intervention d'acteurs secondaires du NHEJ

Bien que Ku et LigD soient les éléments centraux du NHEJ, le mécanisme semble en réalité plus complexe qu'un simple système à deux acteurs. En plus de la protéine LigD, *Mycobacterium* dispose de trois autres ligases : LigA qui est NAD⁺ dépendante et essentielle pour la croissance de la bactérie, ainsi que LigB et LigC deux ligases ATP dépendantes qui ne sont pas indispensables au développement de *Mycobacterium* (Gong *et al.*, 2004). S'il est désormais admis que LigB n'intervient pas dans le NHEJ, LigC peut partiellement compenser une défection de Lig-Dom dans la prise en charge des DSB (Gong *et al.*, 2005 ; Pitcher *et al.*, 2007a ; Bhattarai *et al.*, 2014). De la même manière, les mycobactéries possèdent également deux homologues de Pol-Dom : PolD1 et PolD2. A l'image de Pol-Dom, elles peuvent procéder *in vitro* à l'ajout de nucléotides sur des extrémités franches en l'absence de matrice ou combler des lacunes simple-brin (Zhu *et al.*, 2012) et peuvent jouer un rôle redondant avec Pol-Dom *in vivo* (Bhattarai *et al.*, 2014).

Variabilité du nombre d'acteurs entre les organismes

Même si le mécanisme NHEJ chez *Mycobacterium* est maintenant bien décrit, il est impossible de l'extrapoler aux autres groupes bactériens. Ainsi, le nombre d'acteurs potentiellement impliqués dans le NHEJ peut fluctuer grandement selon les organismes. Un bilan de cette diversité est présenté dans le **tableau I2** pour l'ensemble des procaryotes pour lesquels des études sur le NHEJ ont été menées. Chez *Pseudomonas aeruginosa* il existe un seul exemplaire de *ku* et de *ligD* tandis qu'*Agrobacterium tumefaciens* dispose de trois versions de *ku*, deux de *ligD* et trois de *ligC* (Zhu *et al.*, 2007). Chez les *Sinorhizobium*, bactéries phylogénétiquement proches d'*Agrobacterium tumefaciens*, de multiples homologues de *ku* sont également conservés dans l'ensemble des génomes séquencés (Kobayashi *et al.*, 2008). Enfin, un mécanisme de NHEJ a également été mis en évidence chez l'archée *Methanocella paludicola*. Cette archée possède un gène codant une protéine Ku, ainsi que trois gènes localisés au sein d'un même opéron et codant respectivement une nucléase, une ligase et une polymérase (Bartlett *et al.*, 2013). Bien qu'elles soient codées par des gènes uniques, ces enzymes présentent une forte conservation structurale avec les domaines Nuc-Dom, Pol-Dom et Lig-Dom de leurs homologues bactériens (Bartlett *et al.*, 2015).

Tableau I2 : Diversité des acteurs putatifs du NHEJ chez les procaryotes

Organisme	Groupe	Gènes codant des acteurs putatifs du NHEJ	Références
Bactéries			
<i>Bacillus subtilis</i>	Firmicutes	- <i>ku</i> (<i>ykoV</i>) - <i>ligD</i> (<i>ykoU</i>) : domaine polymérase et domaine ligase	Weller <i>et al.</i> , 2002 ; de Vega, 2013
<i>Mycobacterium tuberculosis</i>	Actinomycètes	- <i>ku</i> - <i>ligD</i> : domaine polymérase, domaine nucléase et domaine ligase - <i>ligC</i>	Gong <i>et al.</i> , 2004 ; Della <i>et al.</i> , 2004 ; Stephanou <i>et al.</i> , 2007 ; Aniukwu <i>et al.</i> 2008 ;
<i>Mycobacterium smegmatis</i>	Actinomycètes	- <i>ku</i> - <i>ligD</i> : domaine polymérase, domaine nucléase et domaine ligase - 2 <i>ligC</i> (<i>ligC1</i> et <i>ligC2</i>)	Gong <i>et al.</i> , 2004 ; Della <i>et al.</i> , 2004 ; Stephanou <i>et al.</i> , 2007 ; Aniukwu <i>et al.</i> 2008 ; Gupta <i>et al.</i> , 2011
<i>Pseudomonas aeruginosa</i>	Protéobactéries	- <i>Ku</i> - <i>ligD</i> : domaine nucléase, domaine ligase et domaine polymérase	Zhu <i>et al.</i> , 2005 ; Zhu <i>et al.</i> , 2006 ; Zhu et Shuman, 2006 ; Zhu et Shuman, 2010
<i>Agrobacterium tumefaciens</i>	Protéobactéries	- 3 <i>ku</i> (<i>ku1</i> à <i>ku3</i>) - 2 <i>ligD</i> : <i>ligD1</i> et <i>ligD2</i> qui possèdent un domaine nucléase, un domaine ligase et un domaine polymérase - 3 <i>ligC</i> (<i>ligC1</i> à <i>ligC3</i>)	Zhu et Shuman, 2007
<i>Sinorhizobium meliloti</i>	Protéobactéries	- 4 copies de <i>ku</i> (<i>sku1</i> à <i>sku4</i>) - 5 copies de <i>ligD</i> : <i>ligD2</i> et <i>ligD4</i> possèdent les trois domaines (nucléase, ligase et polymérase), <i>ligD3</i> possède un domaine ligase et polymérase, <i>ligD1L</i> possède uniquement un domaine nucléase et ligase et <i>ligD1P</i> uniquement un domaine polymérase - <i>ligC</i>	Kobayashi <i>et al.</i> , 2008
Archées			
<i>Methanocella paludicola</i>	Méthanogènes	- <i>ku</i> - un opéron contenant un gène nucléase, polymérase et ligase	Bartlett <i>et al.</i> , 2013 ; Bartlett <i>et al.</i> , 2015

2.3.3 L'alternative end joining chez les bactéries

Chez la bactérie modèle *E. coli* qui ne possède pas de mécanisme NHEJ, une voie de réparation des DSB par recombinaison illégitime alternative à la RH a été mise en évidence (Chayot *et al.*, 2009). Elle se caractérise par des délétions bidirectionnelles et une ligation médiée par microhomologies. Les bactéries du genre *Mycobacterium* dont le NHEJ a largement été décrit dans les paragraphes précédents disposent également d'un *alternative end-joining* (A-EJ) (Aniukwu *et al.*, 2008). Cependant, contrairement au mécanisme découvert chez *E. coli*, l'A-EJ de *Mycobacterium* est hautement fidèle et ne concerne que les extrémités 3' sortantes (Aniukwu *et al.*, 2008). A ce jour, peu d'informations sont disponibles sur l'A-EJ bactérien puisque ces deux études sont les seules traitant d'un tel mécanisme.

2.4 Etat de l'art des systèmes de recombinaison chez les *Streptomyces*

2.4.1 La recombinaison homologue

Comme pour l'ensemble des bactéries, la RH chez les *Streptomyces* est dépendante de la protéine RecA. Bien que le gène *recA* ne soit pas essentiel à la survie de la bactérie, sa délétion entraîne une inactivation de la RH (Vierling *et al.*, 2001 ; Huang *et al.*, 2006). La recherche de complexes hélicase-nucléase homologues de *recBCD* et *addAB* qui sont impliqués dans le traitement des extrémités de la DSB s'est avérée infructueuse dans les génomes de *S. coelicolor* et *S. avermitilis* (Rocha *et al.*, 2005). En revanche, les gènes *adnA* et *adnB* sont bien présent chez les *Streptomyces*. Ils forment un opéron et sont indispensables à la survie de l'organisme (Zhang *et al.*, 2014). Des gènes homologues de *ruvABC* et *recG* connus pour leur implication dans la migration et la résolution des jonctions de Holliday ont été détectés chez *S. coelicolor* et *S. avermitilis* (Rocha *et al.*, 2005), mais aucune étude expérimentale n'a confirmé leur rôle chez les *Streptomyces*.

2.4.2 La recombinaison illégitime

Au travers des différentes études qui ont eu pour objet l'analyse de souches mutantes ayant subi des réarrangements génomiques chez différentes espèces de *Streptomyces*, l'existence d'un mécanisme de recombinaison illégitime est suspectée depuis plusieurs années. Cependant, peu d'informations concernant les acteurs impliqués dans ce système sont disponibles à ce jour. Chez *S. avermitilis*, une équipe de chercheurs s'est focalisée sur la construction de mutants délétés d'une ou des deux copies de *ku* dont dispose le

chromosome (Zhang *et al.*, 2012). Par rapport à une souche sauvage, ces mutants présentent une sensibilité accrue à l'éthyl-methanosulfonate mais pas aux UV ni à la bléomycine. Plus récemment, une étude visant à développer une technique de mutagenèse ciblée en utilisant le système CRISPR-Cas9 a permis de mettre en évidence des délétions de plusieurs kilobases dans le locus cible en absence de matrice pour la RH (Tong *et al.*, 2015). En induisant chez *S. coelicolor* l'expression d'un gène *ligD* complet (codant les trois domaines) issu d'une bactérie du genre *Nocardia*, la taille des délétions observées diminue. Les auteurs ont alors conclu que l'ajout du gène *ligD* a permis de restaurer un NHEJ fonctionnel, qui ne le serait pas dans la situation d'origine en raison de la déficience en *ligD* (Tong *et al.*, 2015). Au sein du laboratoire, des analyses bio-informatiques initiées sur des génomes de *Streptomyces* ont permis de mettre en évidence la présence d'homologues de *ku* dont le nombre et la localisation varient en fonction des espèces. Bien que le gène *ligD* tel qu'il existe chez *Mycobacterium* ne soit pas présent chez *Streptomyces*, on y retrouve en revanche des gènes codant des domaines ligases, polymérases et parfois nucléases de façon individualisée, dont le nombre varie en fonction des espèces. S'il est désormais admis que de nombreux gènes NHEJ-like existent chez les *Streptomyces* ni la démonstration de la fonctionnalité du mécanisme et la nature des acteurs qui le composent, ni l'impact d'un tel système sur la plasticité du génome n'ont été étudiés.

OBJECTIFS DE LA THESE

Bien que la plasticité du génome des *Streptomyces* soit avérée par les études de génomique comparative comme par l'analyse des réarrangements spontanés, les mécanismes de recombinaison et la prise en charge des cassures double-brin le long du génome suspectée d'être à l'origine de la variabilité terminale restent méconnus.

Les objectifs du projet de thèse étaient d'explorer les acteurs de la recombinaison homologue et de la recombinaison illégitime et d'approcher les mécanismes de réparation des DSB par ces deux mécanismes.

Pour cela, après les travaux du laboratoire sur l'acteur majeur de l'étape présynaptique (AdnAB, Zhang *et al.*, 2014), il semblait pertinent de rechercher les acteurs des étapes post-synaptiques de la recombinaison homologue et de tester l'impact de leur inactivation. Le génome de notre modèle *S. ambofaciens*, récemment séquencé (Thibessard *et al.*, 2015 ; Thibessard et Leblond, 2016), a également été analysé afin d'identifier des acteurs potentiels de la recombinaison illégitime, et plus précisément des homologues de gènes du NHEJ. L'identification de ces acteurs potentiels a été poursuivie par une caractérisation du phénotype induit par l'inactivation de ces gènes. Enfin, une stratégie d'induction de DSB ciblées a été mise en place afin d'étudier l'impact de la réparation des DSB sur la structure du génome en fonction de la localisation de la cassure et en fonction du contexte génétique.

Pour la suite de ce manuscrit, les différents résultats seront détaillés en trois parties distinctes. Pour les parties 1 et 2, les résultats sont présentés sous forme d'articles scientifiques. Un résumé en français des résultats exposés, ainsi que certains résultats complémentaires (pour la partie 2) seront également fournis. Pour la partie 3, les résultats seront exposés en français (article en cours de rédaction). Les trois parties s'intituleront comme suit :

- I. Etude des acteurs tardifs de la recombinaison homologue**
- II. Caractérisation des gènes NHEJ-like**
- III. Etude de l'impact de la réparation des DSB sur le génome de *Streptomyces ambofaciens***

Par la suite, ces résultats seront discutés dans une discussion générale et les perspectives à apporter seront détaillées.

CHAPITRE 2 :

-

RESULTATS

I. Etude des acteurs tardifs de la recombinaison homologue

Implication of RuvABC and RecG in Homologous Recombination in *Streptomyces ambofaciens*

Hoff G., Bertrand C., Piotrowski E., Thibessard A., and Leblond P.

Article N°1 : Accepté pour publication dans le journal Research in Microbiology le 04/07/16

Résumé des résultats de l'article :

Dans la partie introductive, le rôle et les caractéristiques des différentes protéines intervenant dans la RH ont été décrits pour les trois bactéries modèles que sont *E. coli*, *B. subtilis* et *Mycobacterium*. Chez les *Streptomyces*, si l'implication de RecA dans la RH a été démontrée à plusieurs reprises (Vierling *et al.*, 2001 ; Huang *et al.*, 2006), l'identité des acteurs intervenant dans les étapes précoces et tardives de la RH est longtemps restée inconnue, faute d'études sur le sujet. Ainsi, après avoir démontré au laboratoire l'essentialité du locus *adnAB*, qui pourrait être responsable du traitement des extrémités d'une DSB lors de l'étape pré-synaptique de la RH (Zhang *et al.*, 2014), nous nous sommes tout naturellement penchés sur l'identification des acteurs de l'étape post-synaptique de la RH. Chez *E. coli* et *B. subtilis*, le résolvasome RuvABC (RuvAB/RecU chez *Bacillus*) et l'hélicase RecG sont connus pour jouer un rôle redondant dans l'étape post-synaptique de la RH (Lloyd et Sharples, 1993 ; Whitby *et al.*, 1993 ; Sanchez *et al.*, 2005). Tandis que RuvAB et RecG catalysent la migration des JH, l'endonuclease RuvC/RecU permet de cliver la structure à quatre brins et de restaurer la conformation double-brin (Connolly *et al.*, 1991 ; Müller et West, 1994 ; Ayora *et al.*, 2004). Ainsi, chez *E. coli*, la délétion de *ruvABC* ou *recG* n'impacte que peu les cellules dans la résistance aux dommages de l'ADN, tandis qu'un mutant $\Delta ruvABC \Delta recG$ y est très sensible (Lloyd et Buckman, 1991). Chez *B. subtilis* la délétion combinée de *ruvAB*, *recU* et *recG* n'est pas viable (Sanchez *et al.*, 2005, 2007).

S'il est connu que les gènes *ruvABC* et *recG* sont présents chez *Streptomyces* (Rocha *et al.*, 2005), aucune information n'est disponible quant à leur conservation au sein des différentes espèces. Ainsi, le premier travail de cette étude a été de rechercher la présence de ces gènes dans les 38 génomes alors séquencés de *Streptomyces*. Nous avons pu constater que les deux loci sont fortement conservés, chaque espèce possédant un orthologue de *ruvABC* et un orthologue de *recG*. Afin de pouvoir étudier leur rôle et implication dans la RH, des souches

mutées pour *ruvABC*, pour *recG* mais également pour les deux *loci* ont été construites chez *S. ambofaciens* ATCC 23877. De la même manière, une souche délétée du gène *recA* a également été construite afin d'être utilisée dans toutes les expériences réalisées comme un témoin déficient pour la RH. Afin de déterminer l'implication de ces gènes dans la RH, 3 approches ont été mises en œuvre.

Dans une première approche, les spores des différents mutants ont été exposées aux rayons UV afin de tester la réponse de ces souches aux dommages de l'ADN. Les résultats nous révèlent que les mutants $\Delta ruvABC$ et $\Delta recG$ ne sont pas significativement plus sensibles aux UV qu'une souche sauvage tandis qu'un double mutant $\Delta ruvABC \Delta recG$ est deux fois plus affecté que la souche sauvage. Toutefois, ces valeurs restent modestes en comparaison d'un mutant $\Delta recA$ dont le taux de survie après exposition aux UV est 250 fois inférieur à celui de la souche sauvage.

Les deux approches suivantes sont axées directement sur des mesures de taux de recombinaison. Elles se distinguent toutefois par le fait que la première approche est basée sur de la recombinaison conjugative, c'est-à-dire l'intégration d'un fragment d'ADN ou le remplacement d'un locus après transfert d'ADN par conjugaison, tandis que la seconde est basée sur de la recombinaison intra-chromosomique déléant un *locus* grâce à un évènement de RH entre deux séquences homologues placées en répétition directe. L'approche de recombinaison conjugative nous révèle ainsi que l'efficacité de recombinaison après conjugaison diminue de 3 à 5 fois pour les mutants $\Delta ruvABC$ et $\Delta recG$, tandis qu'elle est encore aggravée pour un double mutant $\Delta ruvABC \Delta recG$, à savoir 10 à 15 fois inférieure à celle d'une souche sauvage. Cependant, il est à noter que contrairement à ce qui se déroule dans un contexte $\Delta recA$ où aucun recombinant n'a pu être obtenu, la RH est toujours possible en absence de RuvABC et RecG. En revanche, il existe une différence majeure entre recombinaison conjugative et recombinaison intra-chromosomique puisque pour cette dernière, aucune différence significative n'a pu être constatée entre l'efficacité de recombinaison obtenue pour le mutant $\Delta ruvABC \Delta recG$ et la souche sauvage.

Ces résultats nous apportent des informations précieuses sur l'importance de RuvABC et RecG lors de l'étape post-synaptique de la RH chez les *Streptomyces*. En effet, comparés à l'impact qu'entraîne la délétion de ces mêmes acteurs chez d'autres bactéries telles qu'*E. coli* et *B. subtilis*, les résultats obtenus chez *S. ambofaciens* suggèrent que des protéines alternatives seraient également capables d'assurer cette fonction. La différence qui existe entre l'approche conjugative et intra-chromosomique permet également de

supposer que la présence d'ADN exogène pourrait avoir un impact sur l'expression de ces gènes et ainsi favoriser leur intervention plutôt que celle de protéines alternatives.



Research in Microbiology xx (2016) 1–10



www.elsevier.com/locate/resmic

Original article

Implication of RuvABC and RecG in homologous recombination in *Streptomyces ambofaciens*

Grégory Hoff^{a,b}, Claire Bertrand^{a,b}, Emilie Piotrowski^{a,b}, Annabelle Thibessard^{a,b,**},
Pierre Leblond^{a,b,*}

^a Université de Lorraine, Dynamique des Génomes et Adaptation Microbienne, UMR 1128, Vandœuvre-lès-Nancy, F-54506, France

^b INRA, Dynamique des Génomes et Adaptation Microbienne, UMR 1128, Vandœuvre-lès-Nancy, F-54506, France

Received 2 June 2016; accepted 4 July 2016

Available online ■ ■ ■

Abstract

Most bacterial organisms rely on homologous recombination to repair DNA double-strand breaks and for the post-replicative repair of DNA single-strand gaps. Homologous recombination can be divided into three steps: (i) a pre-synaptic step in which the DNA 3'-OH ends are processed, (ii) a *recA*-dependent synaptic step allowing the invasion of an intact copy and the formation of Holliday junctions, and (iii) a post-synaptic step consisting of migration and resolution of these junctions. Currently, little is known about factors involved in homologous recombination, especially for the post-synaptic step. In *Escherichia coli*, branch migration and resolution are performed by the RuvABC complex, but could also rely on the RecG helicase in a redundant manner. In this study, we show that *recG* and *ruvABC* are well-conserved among *Streptomyces*. $\Delta ruvABC$, $\Delta recG$ and $\Delta ruvABC \Delta recG$ mutant strains were constructed. $\Delta ruvABC \Delta recG$ is only slightly affected by exposure to DNA damage (UV). We also show that conjugal recombination decreases in the absence of RuvABC and RecG, but that intrachromosomal recombination is not affected. These data suggest that RuvABC and RecG are indeed involved in homologous recombination in *Streptomyces ambofaciens* and that alternative factors are able to take over Holliday junction in *Streptomyces*.

© 2016 Institut Pasteur. Published by Elsevier Masson SAS. All rights reserved.

Keywords: *Streptomyces*; Homologous recombination; DNA damage; Resolvase

1. Introduction

Among all DNA damage that cells have to cope with, double-strand breaks (DSB) are the most deleterious. They can be induced by various exogenous sources such as chemicals, natural products like mitomycin C or physical forces like desiccation or exposure to ionizing and UV radiations

[1–3]. They can also be triggered by collapse of the replication fork after the replication machinery encounters a single-strand nick [4,5] or by the effect of endogenous genotoxic components such as reactive oxygen species (ROS) [6]. In the absence of repair, DSBs lead to cell death, but in the course of evolution, two major means were devised to repair DSBs: homologous recombination (HR) and illegitimate recombination, including the error-prone non-homologous end joining pathway (NHEJ). In contrast to NHEJ, which can directly ligate the free DNA ends after a facultative processing step, HR needs an intact homologous template to repair the break. DNA repair through HR can be described as a sequence of three stages: the pre-synaptic, synaptic and post-synaptic steps. In bacteria, HR was previously well-documented, especially in some model organisms like *Escherichia coli* or *Bacillus subtilis* [7–15]. In the pre-

* Corresponding author. Université de Lorraine, Dynamique des Génomes et Adaptation Microbienne, UMR 1128, Vandœuvre-lès-Nancy, F-54506, France.

** Corresponding author. Université de Lorraine, Dynamique des Génomes et Adaptation Microbienne, UMR 1128, Vandœuvre-lès-Nancy, F-54506, France.

E-mail addresses: gregory.hoff@univ-lorraine.fr (G. Hoff), claire.bertrand@univ-lorraine.fr (C. Bertrand), emilie.piotrowski@univ-lorraine.fr (E. Piotrowski), annabelle.thibessard@univ-lorraine.fr (A. Thibessard), pierre.leblond@univ-lorraine.fr (P. Leblond).

<http://dx.doi.org/10.1016/j.resmic.2016.07.003>

0923-2508/© 2016 Institut Pasteur. Published by Elsevier Masson SAS. All rights reserved.

Please cite this article in press as: Hoff G, et al., Implication of RuvABC and RecG in homologous recombination in *Streptomyces ambofaciens*, Research in Microbiology (2016), <http://dx.doi.org/10.1016/j.resmic.2016.07.003>

synaptic step, a helicase–nuclease complex binds to the free DNA end and begins to unwind the double-strand molecule while extensively degrading both strands until it reaches specific sequences named *chi*. The main factors involved in this step and the nature of the *chi* sequences can vary among organisms [11,16,17]. The confrontation between the *chi* site and the helicase–nuclease complex causes important changes in the conformation of the latter complex. It leads to the alteration of the 3′-5′ exonuclease activity and therefore to the formation of single-strand 3′ tails on which the RecA protein is loaded. The newly formed RecA nucleofilament promotes homologous pairing and strand invasion during the synaptic step [9,14]. Finally, the post-synaptic step involves the formation, migration and resolution of a four-way DNA intermediate named the Holliday junction (HJ) and, at least in *E. coli*, PriA-dependent loading of the replicative helicase on the post-synaptic DSB repair intermediate called a D-loop leads to the assembly of a replication fork and couples DSB repair with replication restart [18–20].

A high number of studies concerning the post-synaptic step of HR in bacteria were undertaken in the past decades, but involvement and importance of each factor remain elusive. In *E. coli*, the proteins RuvA, RuvB and RuvC are known to form a tripartite complex able to displace and resolve an HJ [21]. In the presence of ATP, RuvA binds DNA and, together with RuvB, drives the branch migration of the HJ, leading to the formation of a DNA heteroduplex [22,23]. In concert with RuvAB, RuvC resolves the newly formed structure by a dual-strand incision [24]. The nicked DNA duplexes can then split up and be restored by DNA ligases [25,26]. HJ migration can also be performed by the RecG helicase [27–29]. Although the role of RuvAB in branch migration and that of RuvABC in resolution of HJs is firmly established, the role of RecG remains more elusive [30]. Recently, RecG was proposed to act on the D-loop formed by RecA filament invasion, rather than on the HJ, and to promote the correct loading of PriA, allowing efficient replication restart and the formation of a viable recombinant molecule [31]. The additive defects conferred by *ruvABC* and *recG* mutations would then result from the fact that HJ branch migration and replication restart from the D-loop are two ways of stabilizing the recombination intermediate. Even though *ruvAB* and *recG* are highly conserved among bacterial genomes [32], the impact of their loss on cell metabolism is variable between organisms. *E. coli* devoid of RuvABC or RecG displays weak sensitivity to DNA damage and a slight decrease in HR efficiency. In contrast, a double $\Delta ruvABC \Delta recG$ mutant strain shows marked aggravation of both phenotypes, suggesting that the two pathways achieve redundant roles [27]. In *B. subtilis*, the single deletion of either *ruvAB*, or *recG* or *recU* (a functional homologue of *ruvC*) is highly deleterious for cell survival upon DNA damage, whereas the different combinations ($\Delta ruvAB \Delta recU$, $\Delta ruvAB \Delta recG$ and $\Delta recG \Delta recU$) of double mutation are lethal [12,33].

Streptomyces are Gram-positive soil bacteria belonging to the Actinomycetales order. They are characterized by a complex lifestyle in which a vegetative mycelium gives rise to an

aerial mycelium differentiating in spore chains. They are also known to produce a large panel of secondary metabolites. In contrast to most bacterial phyla, they possess a linear chromosomal DNA with a typical replication origin located approximately at the middle of the chromosome [34,35] and terminal proteins attached covalently to the 5′ ends of the DNA [36]. The genome of the *Streptomyces* species presents a high plasticity characterized by frequent rearrangements resulting from both homologous recombination triggering chromosomal arm replacement [37,38], and illegitimate recombination leading, for instance, to chromosomal circularization (for review [39,40]). Extensive DNA amplifications are also frequently encountered on rearranged chromosomes [41], resulting from an unknown mechanism involving homologous recombination [42].

At the moment, little is known about HR factors in *Streptomyces*. While there is no homologue for either *recBCD* of *E. coli* or *addAB* of *B. subtilis* in the *Streptomyces* genomes thus far sequenced, these bacteria possess an *adnAB* locus [43], as described in other actinomycetes [17]. This locus encodes a helicase–nuclease complex meant to play the role of *recBCD* and *addAB* in *E. coli* and *B. Subtilis*, respectively, which is to process the DNA ends of a DSB and initiate the recombination repair pathway. In contrast to *Mycobacterium tuberculosis*, the deletion of this locus was not possible in *Streptomyces ambofaciens*, strongly suggesting that the encoded function was essential [43]. On the other hand, a mutant deficient for *recA* is viable in *Streptomyces* [44]. $\Delta recA$ strains are particularly sensitive to UV and mitomycin C exposure and are unable to perform HR in conjugational recombination [44]. Concerning the post-synaptic step of HR, *in silico* analyses revealed the presence of *ruvABC* and *recG* homologues in *Streptomyces coelicolor* and *Streptomyces avermitilis* [32]. Assessing their role in *S. ambofaciens* HR is the purpose of this study.

2. Materials and methods

2.1. Bacterial strains and culture conditions

All strains used in this study are listed in Table 1. The DH5 α *E. coli* strain is used as a conservation host for BACs and plasmids. The ET12567 non-methylating strain containing the mobilizing pUZ8002 plasmid was used as donor for intergeneric conjugation with *S. ambofaciens*. All *E. coli* strains were grown in Luria-Bertani (LB) medium at 37 °C, except for the BW25113 pKD20 thermosensitive strain used for PCR targeting, which was grown at 30 °C. *S. ambofaciens* ATCC 23877 was our reference strain from which the mutant strains derived. All *S. ambofaciens* strains were grown at 30 °C on soya flour mannitol (SFM) plates, except for UV assays, which were performed on solid Hickey Tresner (HT) medium. Liquid HT medium was used for growth of mycelium before DNA extraction. When necessary, antibiotics were added to the medium at a concentration of 50 $\mu\text{g}/\mu\text{L}$ for apramycin and kanamycin, and 25 $\mu\text{g}/\mu\text{L}$ for nalidixic acid and chloramphenicol. For intrachromosomal recombination

Table 1
Strains, plasmids and BACs used in this study.

Strains, plasmids and BACs	Characteristics	Sources
<i>Streptomyces</i>		
<i>S. ambofaciens</i> ATCC 23877	Used as a wild-type strain	[45]
<i>S. ambofaciens</i> ATCC 23877 Δ <i>ruvABC</i>	Deletion of the <i>ruvABC</i> locus	This study
<i>S. ambofaciens</i> ATCC 23877 Δ <i>recG</i>	Deletion of the <i>recG</i> locus	This study
<i>S. ambofaciens</i> ATCC 23877 Δ <i>recG</i> Δ <i>ruvABC</i>	Deletion of the <i>recG</i> and the <i>ruvABC</i> loci	This study
<i>S. ambofaciens</i> ATCC 23877 Δ <i>recA</i>	Deletion of the <i>recA</i> locus	This study
<i>E. coli</i>		
DH5 α	<i>F</i> ⁻ ϕ dlacZ Δ M15 <i>endA1 supE44 thi-1 recA1 relA1 gyrA96 deoR nupG</i> Δ (<i>lacZYA-argF</i>)U169 λ - <i>hsdR17</i> (<i>r_gm_k⁺</i>) <i>phoA relA1 dam-13::Tn9 dcm cat tet hsdM</i>	[46]
ET12567/pUZ8002	<i>hsdR zjj-201::Tn10 tra neo RP4</i>	[47]
BW25113/pKD20	K12 derivative Δ <i>araBAD</i> , Δ <i>rhaBAD</i> λ -RED (<i>gam, bet, exo</i>), <i>bla, araC, rep101^{ts}</i>	[48]
Plasmids/BACs ^a		
BAA8H12 Δ SAM1569-SAM1571	Δ <i>chl::kan</i> Δ <i>ruvABC::aac3(IV)-oriT</i>	This study
BAB10F9 Δ SAM5320	Δ <i>chl::kan</i> Δ <i>recG::aac3(IV)-oriT</i>	This study
BAB24H12 Δ SAM5320	Δ <i>chl::kan</i> Δ <i>recG::aac3(IV)-oriT</i>	This study
BAB20B5 Δ SAM5473	Δ <i>chl::kan</i> Δ <i>recA::aac3(IV)-oriT</i>	This study
BAB13C7 Δ SAM1172	Δ <i>chl::kan</i> Δ <i>cebR::aac3(IV)-oriT</i>	This study
pDYN6902	<i>E. coli</i> to <i>Streptomyces</i> conjugative and integrative shuttle vector	[49]

^a All original BACs arise from [50].

assays, bacterial lawns were picked using a 2 cm diameter cookie-cutter and spores were harvested after being thoroughly vortexed (3 times 20 s). For conjugational recombination assays, the number of *Streptomyces* spores was calibrated to 10⁸ UFC and added to 2 · 10⁸ UFC of the *E. coli* donor culture.

2.2. Mutant strain construction

Mutant strains were constructed by PCR targeting as described by Gust et al., 2003 [51]. To summarize, a recombinant BAC containing the locus of interest was transformed in the highly recombinogenic *E. coli* BW25113/pKD20 strain. The target gene was replaced by an apramycin resistance disruption cassette [52] in which an *oriT* sequence as introduced. The modified BAC was then transferred to the *S. ambofaciens* ATCC 23877 strain using the ET12567/pUZ8002 donor host. Double-crossover (CO) events leading to replacement of the target gene by the cassette were selected by sensitivity to kanamycin (loss of the BAC vector part) along with resistance to apramycin (insertion of the resistance cassette). The disruption cassette flanked by *attL* and *attR* sites was excised in the newly obtained mutant strains after the introduction of pOSV508, a plasmid allowing expression of *int* and *xis* genes from pSAM2 [52]. Excision was verified by PCR after DNA extraction using a low binding sulfate salt method [53]. All the primers used in this experiment are listed in Table 2.

2.3. Ultraviolet (UV) light exposure and viability rates

Serial dilutions of the spore suspension to be tested were spread on HT plates and exposed to 0, 75 and 150 J m⁻² UV

doses using a 254 nm UV light. Plates were immediately protected from light to avoid photoreactivation and then incubated at 30 °C [54]. Surviving colonies were counted after 3 days of growth, and survival rates were calculated in comparison to the same dilutions of unexposed cells of the same culture. For each condition, three biological replicates were carried out.

2.4. Statistical analyses

For UV light exposure as well as for recombination assays, comparisons were achieved pairwise with a *t* test under R and adjusted with the Bonferroni correction [55].

3. Results

3.1. *ruvABC* and *recG* are not essential for *Streptomyces development*

In analysis aimed at identification of orthologs of DNA recombination genes, Rocha et al. [32] revealed the presence of orthologs of *ruvABC* and *recG* in *S. coelicolor* and *S. avermitilis* genomes. To determine whether *ruvABC* and *recG* loci are well-conserved among *Streptomyces* species, we screened the 36 fully sequenced genomes of *Streptomyces* species (tBLASTN algorithm of the genome database of NCBI). Based on the protein sequence of the genes labelled as *ruvA*, *ruvB*, *ruvC* and *recG* in *S. coelicolor*, we found an ortholog for each gene in each genome, with a minimal identity of 69% for *SCO1519*, 82% for *SCO1518*, 68% for *SCO1520* and 73% for *SCO5556* encoding RuvA, RuvB, RuvC and RecG, respectively. The *ruvABC* genes are co-

Table 2
Primers used in this study.

Primer name	Sequence (5'-3')	Characteristics
kanDchl1	AGGTTCCAACCTTTACCCATAATGAAATAAGATCACTACCGCCAGTTCGCCCAATTCTC	Replacement of <i>chl</i> by <i>neo</i>
kanDchl2	TGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGAACCATGAGATCCCCGCGCTGGA	Replacement of <i>chl</i> by <i>neo</i>
ruv_L3	GTGCGCGTACTGGGGGTGGACCCGGGACTGACCCGTCGCGGTGCCTCTTCGTCCCGAAGCAACTTG	<i>ruvABC</i> deletion
ruv_R3	GGACGCGGGCGGGTGAAGGCCAGGTGCGCCACGCGGCCGGGCGCGCTTCGTTCGGGACGAAGAGG	<i>ruvABC</i> deletion
ruv11	CGAGGAATTCGGGTGCGGTACTTCCTTCAACA	Verification of <i>ruvABC</i> deletion
ruv12	GCGAGAAITCCATGGAGAAGGTCGACGCGT	Verification of <i>ruvABC</i> deletion
D_recG_F	TGGCGTGGTGTGCAATGGATCTCGTGCCCGCACTGCGAGTGCCTCTTCGTCCCGAAGCAACTTG	<i>recG</i> deletion
D_recG_R	TCAGCCCTTCTCCAGGTACTGCTCCCTCTCCTCGTCCAGGCGCGCTTCGTTCGGGACGAAGAGG	<i>recG</i> deletion
V_recG_F	CGTGCTCATCGGCGTG	Verification of <i>recG</i> deletion
V_recG_R	TGTGTTCCGTTTCAGTCTGGC	Verification of <i>recG</i> deletion
DrecA_F	GTGAAGCGATCGAATCAAGCAAACCGGGTGGAAACCCATGGTGCCTCTTCGTCCCGAAGCAACTTG	<i>recA</i> deletion
DrecA_R	CCCAGTCCGTTCCGTCGTGTACGGGTACGCTCTTGGTGCAGCGCGCTTCGTTCGGGACGAAGAGG	<i>recA</i> deletion
recA11	CCGGGAATTCGGACGTACGCTCGGAGGTGC	Verification of <i>recA</i> deletion
recA12	TGGGGAATTCGTCTCCTCCGCTCACGCTCG	Verification of <i>recA</i> deletion
SaCebRmut_F	CCGCAGAGTGCGGGTCTCGACGGGAGGCGGAGCCGATGATCGCGCGCGCTTCGTTCGGGACGAA	Deletion of SAM23877_1172
SaCebRmut_R	CCGCACGGCACCCCGTGCCGCCCGCTCCCCACCCCTCAATCTGCCTCTTCGTCCCGAAGCAACT	Deletion of SAM23877_1172

localized and form a potential *ruvCAB* operon. In *S. ambifaciens*, RuvA, RuvB, RuvC and RecG (corresponding to SAM23877_1570, SAM23877_1569, SAM23877_1571 and SAM23877_5320) have a predicted size of, respectively, 201, 357, 233 and 745 amino acids.

We employed the PCR targeting method to construct mutant strains in order to test the involvement of these genes in HR. We generated strains deleted for the *ruvABC* operon or for the *recG* locus, as well as double-mutant strains. In a HR-defective control context, we targeted the *recA* locus. For each genotype, two independent BACs were modified by replacement of the target gene(s) by a resistance cassette and were used to generate two independent mutant strains. All mutant strains were obtained with classical frequencies, showing that none of these genes was essential for growth in laboratory conditions or requested selection for additional suppressor mutations. No specific colony morphology phenotype or sporulation deficiency was observed for the $\Delta ruvABC$ or $\Delta recG$ strains compared to the wild-type (WT) when grown as single colonies along a time course growth of 96 h, corresponding to the entry in sporulation phase for the WT strain

(HT medium, 30 °C). In contrast, the double mutant $\Delta ruvABC \Delta recG$ showed a growth phenotype characterized by a slight pigmentation delay (Fig. 1).

3.2. The double $\Delta ruvABC \Delta recG$ mutation confers sensitivity to UV light exposure

Spores were exposed to UV at 75 J m⁻² and 150 J m⁻² immediately after plating on solid HT medium, and the survival rate was calculated relative to unexposed cells. The WT strain showed cell survival rates of 66.45 ± 4.29% and 25.02 ± 2.18% at UV doses of 75 J m⁻² and 150 J m⁻², respectively (Fig. 2). The $\Delta ruvABC$ and $\Delta recG$ strains showed a sensitivity of 61.93 ± 3.87% and 59.12 ± 4.51% at 75 J m⁻² and 19.93 ± 1.59% and 17.01 ± 2.12% at 150 J m⁻², respectively. Even if though values were slightly lower than for the WT, the difference was not statistically significant. In contrast, with sensitivity rates of 43.68 ± 3.52% at 75 J m⁻² and 11.95 ± 1.66% at 150 J m⁻², the $\Delta ruvABC \Delta recG$ mutant strains were weakly but significantly more sensitive to UV exposure than the WT (at risk $\alpha = 0.05$). This sensitivity was

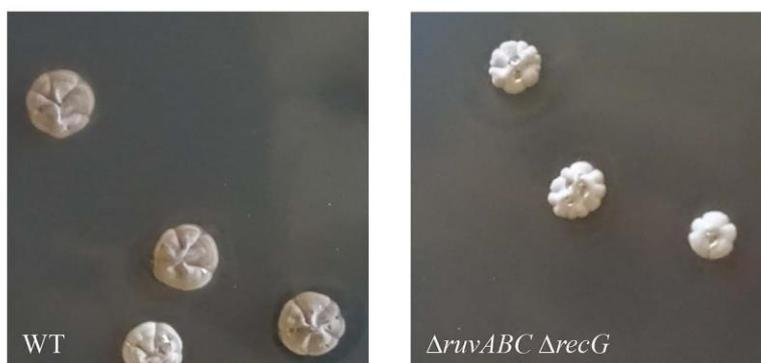


Fig. 1. Phenotype of *S. ambifaciens* ATCC 23877 WT and $\Delta ruvABC \Delta recG$ strains after 96 h growth on HT medium at 30 °C. In comparison to WT, a $\Delta recG \Delta ruvABC$ strain shows a slight pigmentation delay.

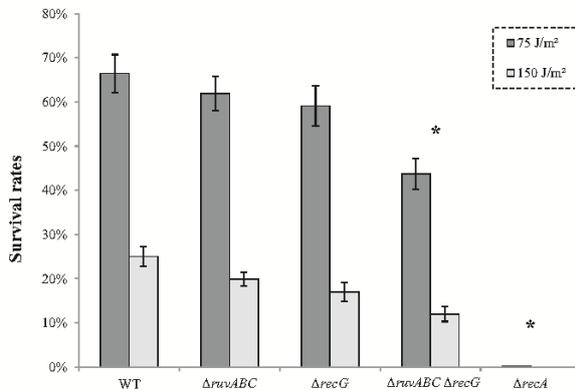


Fig. 2. Survival rates of WT, $\Delta ruvABC$, $\Delta recG$, $\Delta ruvABC \Delta recG$ and $\Delta recA$ strains after UV exposure at 75 J m^{-2} (dark gray) and 150 J m^{-2} (light gray). This experiment was performed in triplicate. The asterisk (*) indicates a significant difference from the WT (*t*-test) at risk $\alpha = 0.05$.

much less pronounced in comparison to the $\Delta recA$ strain, which showed a drastic decrease in resistance rates, with a value of $0.11 \pm 0.08\%$ at the lowest tested dose.

3.3. The double $\Delta ruvABC \Delta recG$ mutant is only affected in formation of conjugational recombinants

To test the ability of the $\Delta recG \Delta ruvABC$ mutant to achieve homologous recombination, we compared the mutant strains and the WT reference for their capacity to form recombinants (Figs. 3 and 4). For that purpose, we targeted a locus not implicated in the DNA recombination processes (namely *SAM23877_1172* encoding the *Streptomyces* cellobiose sensor [56]). In a first approach (Fig. 3), we introduced by intergeneric conjugation (see Materials and methods) the BAC in which *SAM23877_1172* is replaced by an apramycin resistance cassette; then, apramycin resistant (Apra^R) clones were selected for and counted. By this means, we counted all the recombination events, resulting in the integration of the apramycin resistance gene into the chromosome, the most probable being a single cross-over occurring between the insert of the recombinant vector and its homologous counterpart on the chromosome (recombination within the H1 or H2 region; the event within H1 is represented on Fig. 3A). By doing so, we also accounted for the minor proportion of recombinants that directly underwent a double cross-over leading to chromosome allele replacement (recombination within H1 and H2, Fig. 3A).

Apra^R clones resulting from single or double cross-over events were counted in each genetic context and compared to the WT (Fig. 3B). The ratio of Apra^R clones obtained in $\Delta ruvABC$ and $\Delta recG$ contexts was $28.71 \pm 5.16\%$ and $19.82 \pm 3.64\%$, respectively, relative to the WT, showing that conjugational recombination efficiency was reduced on average 3.5-fold and 5-fold in *ruvABC*- and in *recG*-defective mutants respectively. In the homologous recombination-deficient *recA* mutant, no Apra^R clone could be observed in

the same experimental conditions. When both loci are deleted, the recombination efficiency was further reduced to $7.4 \pm 2.38\%$ of the WT reference. These results demonstrated that *ruvABC* and *recG* are involved in conjugational recombination. Since the recombination-deficient phenotype is aggravated in double $\Delta ruvABC \Delta recG$ mutant strains compared to single mutants, we concluded that RuvABC and RecG may be redundant for formation of recombinants. However, it should be noted that, in the absence of both RuvABC and RecG, HR is still effective, suggesting that alternative proteins can catalyze migration and resolution of Holliday junctions in this double mutant.

Since our observations followed an intergeneric *E. coli/ Streptomyces* conjugation, differences in the conjugational capacities of the WT and the $\Delta ruvABC \Delta recG$ genetic backgrounds may be responsible for the decrease in recombinants. To rule out this possibility, we set up conjugation experiments using the conjugative and integrative plasmid pDYN6902 (derived from pIJ6902); its chromosomal integration occurs by site-specific recombination and is therefore independent from HR [49]. This vector shares the same replication origin as pBeloBAC11 used to construct the BAC library [50], and is thus expected to have an equivalent copy number in the donor *E. coli* strain. Since its integration is very efficient, we monitored the conjugation frequency of the conjugative and integrative pDYN6902 plasmid in the mutant strain and compared it to the WT. There was no significant difference between the two strains, with conjugational rates of $1.2 \cdot 10^{-4}$ and $9.7 \cdot 10^{-5}$ for the WT and $\Delta ruvABC \Delta recG$ strains, respectively.

We further set up a second experiment to compare the frequency of homologous recombination in the different mutants (Fig. 4). In this approach, we assayed the frequency of gene replacement (*SAM23877_1172* by the apramycin resistance cassette). This step corresponds to formation of a crossing-over occurring between the long tandem repeats resulting from integration of the recombinant BAC into the genome. The second cross-over event removes vector DNA, including the kanamycin resistance gene, leading to a $\text{Apra}^R \text{Kan}^S$ phenotype from an $\text{Apra}^R \text{Kan}^R$ background. We used an experimental approach inspired by the classical Luria and Delbrück fluctuation test [57]. The idea was to test the occurrence of the second cross-over in independent samples of WT and $\Delta ruvABC \Delta recG$ bacterial populations from clones that have integrated the recombinant vector by a single cross-over. Twenty samples of bacterial lawns (picked using a cookie-cutter, see Materials and methods section) and thus corresponding to equal subpopulations were harvested for WT and $\Delta ruvABC \Delta recG$ strains in which the first cross-over happened ($\text{Apra}^R \text{Kan}^R$ transconjugants). Spore suspensions from these lawn samples were titrated and grown on HT medium supplemented with only apramycin to retain the presence of the resistance cassette, but without counter-selecting the formation of the second crossing-over. Plates were further replica-plated on HT medium supplemented with both apramycin and kanamycin to determine the frequency of $\text{Apra}^R \text{Kan}^S$ clones among the Apra^R .

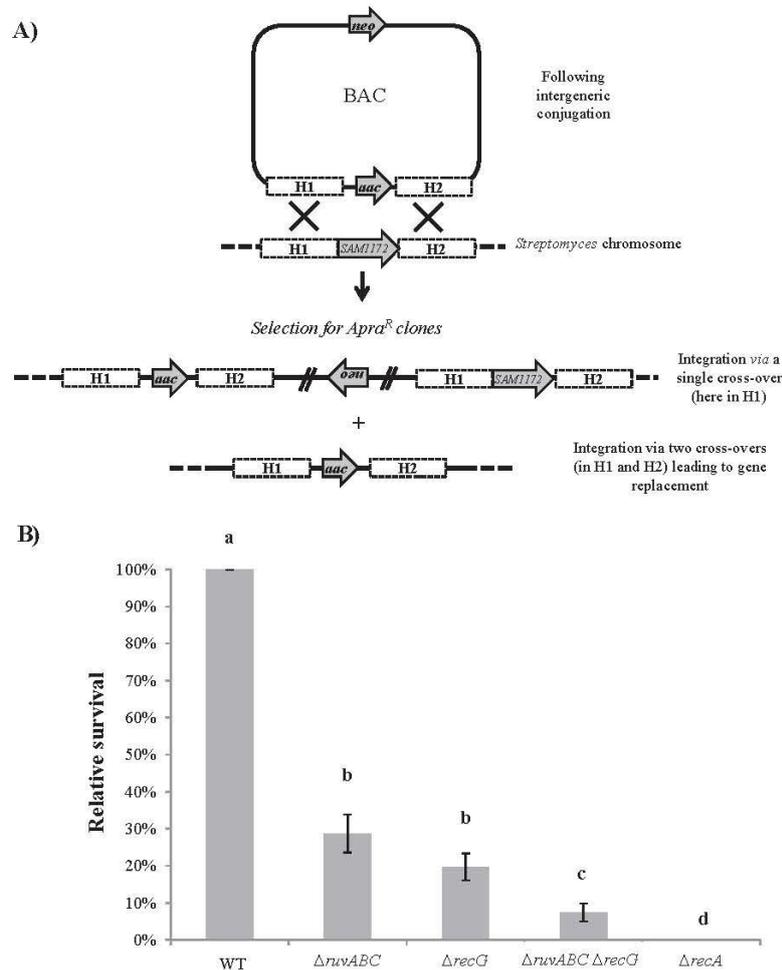


Fig. 3. Conjugational recombination efficiency in *S. ambofaciens*. **A)** Representation of single or double cross-over events occurring between the *Streptomyces* chromosome and a homologous DNA transferred by intergeneric conjugation. The *neo* gene confers resistance to kanamycin and *aac* stands for *aac(3)IV* which encodes the aminoglycoside 3-N-acetyltransferase and confers resistance to apramycin. **B)** Proportion of single and double cross-over events represented by the proportion of Apra^R clones obtained after conjugation in $\Delta ruvABC$, $\Delta recG$, $\Delta ruvABC \Delta recG$, $\Delta recA$ and WT strains. Values of the mutant strains were normalized to that of WT. The letters a, b, c and d indicate the different statistical groups (*t*-test) at risk $\alpha = 0.05$.

For both genetic contexts, at least one Apra^R Kan^S clone was counted in each of the 20 samples of population, revealing that the second cross-over could occur at an equivalent frequency (Fig. 4B). Furthermore, Apra^R Kan^S clones were observed at frequencies which were not significantly different (risk $\alpha = 0.05$): $13.38 \pm 4.95\%$ and $11.67 \pm 4.98\%$ for the WT and the $\Delta ruvABC \Delta recG$ respectively. According to Luria and Delbrück, this reveals that intra-chromosomal homologous recombination is not significantly affected in the $\Delta ruvABC \Delta recG$ double-mutant strain compared to WT.

These data show that the deficiency in RuvABC and RecG does not significantly affect intrachromosomal homologous recombination in *S. ambofaciens*, while it decreases the capacity to form recombinants after conjugational transfer.

4. Discussion

4.1. Alternative homologous recombination post-synaptic proteins functioning in *Streptomyces*

RuvABC resolvosome and RecG branch migration enzyme are assumed to act as alternative proteins for the post-synaptic step of homologous recombination in *E. coli* [8,58]. While RuvAB and RecG are able to perform branch migration after HJ formation, the endonuclease RuvC can resolve the four-way structure and restore the two double-strand homologs [59]. In addition to branch migration of HJs, in vitro RecG can act on a variety of structures and promote a variety of reactions, so that, to date, its role in vivo remains undetermined [30]. In this study, we found a high level of conservation of

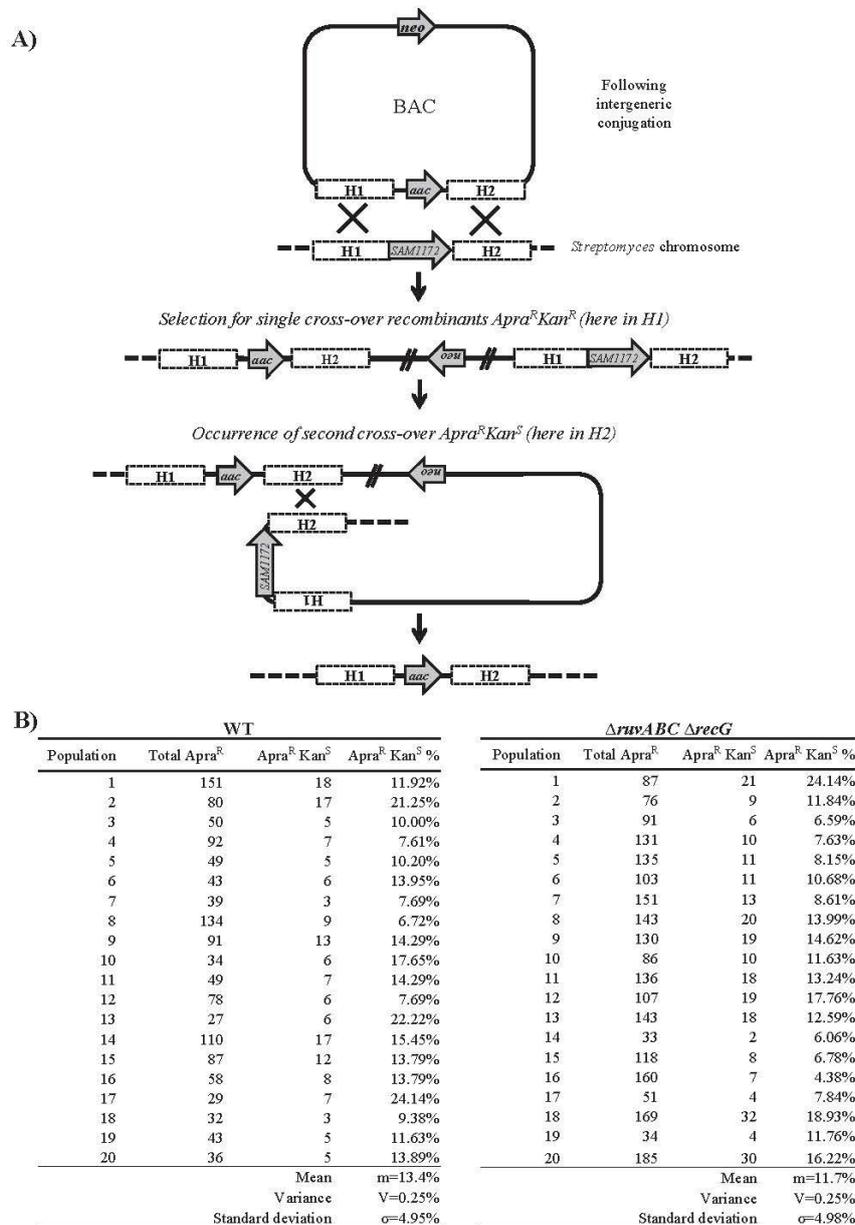


Fig. 4. Intrachromosomal recombination efficiency in *S. ambofaciens*. A) Representation of a second cross-over event occurring within the chromosome after a single cross-over event. B) Proportion of second cross-over events represented by the proportion of Apra^R Kan^S clones in the 20 populations obtained after a single cross-over event in WT and $\Delta ruvABC \Delta recG$ backgrounds.

ruvABC and *recG* in all *Streptomyces* genomes sequenced so far. We studied the involvement of both in HR in our bacterial model *S. ambofaciens*. For this purpose, we measured the sensitivity to UV irradiation (a DNA damaging agent), we quantified homologous recombination frequency after conjugational transfer and homologous recombination between tandemly repeated sequences on the chromosome. $\Delta ruvABC \Delta recG$ is more sensitive to UV irradiation than wild-type cells, and in a $\Delta ruvABC \Delta recG$ mutant strain, we detected a 7.4-fold

decrease in HR efficiency in conjugational recombination. In contrast, no difference in intra-chromosomal recombination was observed between the mutant and the WT strain. The latter result shows that proteins may compensate for the deficiencies of RuvABC resolvase and the RecG enzyme. An open question is how to explain the significant difference in HR measured by conjugation and by intramolecular recombination assay. In conjugational HR, the substrate DNA (exogenote) is transferred from *E. coli* to *Streptomyces* as a

single-strand DNA. In *E. coli*, it is well-known that persistent single-strand DNA leads to the induction of the SOS system [60]. This favors interspecies recombination through the increasing expression of *recA* and *ruvAB* [60]. Hence, we can extrapolate and hypothesize that the single-stranded transfer of the BAC DNA from *E. coli* to *Streptomyces* also stimulates expression of these two recombination enzymes. Similarly, UV irradiation is known to induce the SOS response. Increased expression of RecA may stimulate the first steps of HR, increasing the need for the enzymes that act at later steps. In contrast, intrachromosomal recombination would not be promoted by SOS-induction of RecA, and the RuvABC/RecG defect could be compensated by the basic expression of alternative resolvases.

4.2. Migration and resolution of HJ, a high redundant function?

In the bacterial models *E. coli* and *B. subtilis*, RuvABC/RecU and RecG are known to be the main factors of the HR post-synaptic steps [7,12,33,58]. Unexpectedly, the simultaneous inactivation of *ruvABC* and *recG* confers a mild phenotype in response to DNA damage and in HR efficiency in *S. ambifaciens*, suggesting the existence of alternative ways of processing post-synaptic molecules. In *E. coli*, when the late stages of HR are blocked (e.g. through *ruvABC* mutations), the loss of *uvrD*, known to be involved in mismatch and nucleotide excision repair, is lethal [61]. In addition, *E. coli* UvrD was recently shown to be able to unwind HJ in vitro and it could therefore have a redundant role with RuvAB and RecG [62]. In the *S. ambifaciens* genome, two *uvrD* homologues can be detected. Another candidate for branch migration could be a *recA* paralog known as *sms* or *radA* that is well conserved among all bacterial genomes sequenced thus far [63]. In *E. coli*, measurements of conjugational recombination suggest a redundant function for *radA/sms* with the other branch migration factors [64]. Recently, a study showed that RadA/Sms is able to perform branch migration, but it is suspected to act differently than RuvAB and RecG, since RadA/Sms displays no helicase activity [65].

For HJ resolution activity, an alternative candidate can also be suggested. Aravind et al. [66] identified the YqgF family as sharing homologous and structural relationships with the RuvC family, suggesting that its members, predicted resolvases, could function as an alternative to RuvC. While preparing this manuscript, Nautiyal et al. [67] reported in vitro analyses carried out on RuvX (the mycobacterial member of the YqgF family). They revealed that this protein has affinity for HJs and displays HJ cleavage activity. *Mycobacteria* and *Streptomyces* both belong to the actinobacteria and an ortholog to RuvX can be identified in the *Streptomyces* genomes. Altogether, these data support the idea that RuvX is a serious candidate for resolving HJs as an alternative to RuvC in *S. ambifaciens*. It would be interesting to investigate the involvement of the aforementioned putative alternative factors in the recombination process in *Streptomyces*.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

This work was supported by the French National Research Agency program Streptoflux (ANR Blanc 0096_01), by ANR through the Laboratory of Excellence ARBRE (ANR-12-LABXARBRE-01) and by the Région Lorraine. We are grateful to Dr. Bénédicte Michel (CNRS, Gif-sur-Yvette, France) for helpful discussions and critical reading of the manuscript.

References

- [1] Bonura T, Smith KC. Enzymatic production of deoxyribonucleic acid double-strand breaks after ultraviolet irradiation of *Escherichia coli* K-12. *J Bacteriol* 1975;121:511–7.
- [2] Ward JF. Nature of lesions formed by ionizing radiation. DNA damage and repair. Springer; 1998. p. 65–84.
- [3] Vignard J, Mirey G, Salles B. Ionizing-radiation induced DNA double-strand breaks: a direct and indirect lighting up. *Radiother Oncol* 2013;108:362–9.
- [4] Michel B, Ehrlich SD, Uzest M. DNA double-strand breaks caused by replication arrest. *EMBO J* 1997;16:430–8.
- [5] Mehta A, Haber JE. Sources of DNA double-strand breaks and models of recombinational DNA repair. *Cold Spring Harb Perspect Biol* 2014;6:a016428–a016428.
- [6] Li Y, Trush MA, Yager JD. DNA damage caused by reactive oxygen species originating from a copper-dependent oxidation of the 2-hydroxy catechol of estradiol. *Carcinogenesis* 1994;15:1421–7.
- [7] Taylor AF. Movement and resolution of Holliday junctions by enzymes from *E. coli*. *Cell* 1992;69:1063–5.
- [8] Lloyd RG, Sharples GJ. Processing of recombination intermediates by the RecG and RuvAB proteins of *Escherichia coli*. *Nucleic Acids Res* 1993;21:1719–25.
- [9] Kowalczykowski SC, Eggleston AK. Homologous pairing and DNA strand exchange proteins. *Annu Rev Biochem* 1994;63:991–1043.
- [10] Myers RS, Stahl FW. Chi and the RecBCD enzyme of *Escherichia coli*. *Annu Rev Genet* 1994;28:49–70.
- [11] Chedin F, Ehrlich SD, Kowalczykowski SC. The *Bacillus subtilis* AddAB helicase/nuclease is regulated by its cognate Chi sequence in vitro. *J Mol Biol* 2004;298:7–20.
- [12] Sanchez H, Kidane D, Reed P, Curtis FA, Castillo Cozar M, Graumann PL, et al. The RuvAB branch migration translocase and RecU Holliday junction resolvase are required for double-stranded DNA break repair in *Bacillus subtilis*. *Genetics* 2005;171:873–83.
- [13] Chedin F, Handa N, Dillingham MS, Kowalczykowski SC. The AddAB helicase/nuclease forms a stable complex with its cognate sequence during translocation. *J Biol Chem* 2006;281:18610–7.
- [14] Cox MM. Motoring along with the bacterial RecA protein. *Nat Rev Mol Cell Biol* 2007;8:127–38.
- [15] Dillingham MS, Kowalczykowski SC. RecBCD enzyme and the repair of double-stranded DNA breaks. *Microbiol Mol Biol Rev* 2008;72:642–71.
- [16] Bianco PR, Kowalczykowski SC. The recombination hotspot Chi is recognized by the translocating RecBCD enzyme as the single-strand of DNA containing the sequence 5'-GCTGGTGG-3'. *Proc Natl Acad Sci U S A* 1997;94:6706–11.
- [17] Sinha KM, Unciuleac M-C, Glickman MS, Shuman S. AdnAB: a new DSB-resecting motor-nuclease from mycobacteria. *Genes Dev* 2009;23:1423–37.
- [18] Kogoma T, Cadwell GW, Barnard KG, Asai T. The DNA replication priming protein, PriA, is required for homologous recombination and double-strand break repair. *J Bacteriol* 1996;178:1258–64.

- [19] Sandler SJ, Samra HS, Clark AJ. Differential suppression of *priA2*: Kan phenotypes in *Escherichia coli* K-12 by mutations in *priA*, *lexA*, and *dnaC*. *Genetics* 1996;143:5–13.
- [20] Sandler SJ, Marians KJ. Role of PriA in replication fork reactivation in *Escherichia coli*. *J Bacteriol* 2000;182:9–13.
- [21] Grove JJ, Harris L, Buckman C, Lloyd RG. DNA double-strand break repair and crossing over mediated by RuvABC resolvase and RecG translocase. *DNA Repair* 2008;7:1517–30.
- [22] Tsaneva IR, Müller B, West SC. ATP-dependent branch migration of Holliday junctions promoted by the RuvA and RuvB proteins of *E. coli*. *Cell* 1992;69:1171–80.
- [23] Parsons CA, Tsaneva I, Lloyd RG, West SC. Interaction of *Escherichia coli* RuvA and RuvB proteins with synthetic Holliday junctions. *Proc Natl Acad Sci U S A* 1992;89:5452–6.
- [24] Connolly B, Parsons CA, Benson FE, Dunderdale HJ, Sharples GJ, Lloyd RG, et al. Resolution of Holliday junctions *in vitro* requires the *Escherichia coli* *ruvC* gene product. *Proc Natl Acad Sci U S A* 1991;88:6063–7.
- [25] Dunderdale HJ, Benson FE, Parsons CA, Sharples GJ, Lloyd RG, West SC. Formation and resolution of recombination intermediates by *E. coli* RecA and RuvC proteins. *Nature* 1991;354:506–10.
- [26] Iwasaki H, Takahagi M, Shiba T, Nakata A, Shinagawa H. *Escherichia coli* RuvC protein is an endonuclease that resolves the Holliday structure. *EMBO J* 1991;10:4381.
- [27] Lloyd RG, Buckman C. Genetic analysis of the *recG* locus of *Escherichia coli* K-12 and of its role in recombination and DNA repair. *J Bacteriol* 1991;173:1004–11.
- [28] Lloyd RG, Sharples GJ. Dissociation of synthetic Holliday junctions by *E. coli* RecG protein. *EMBO J* 1993;12:17.
- [29] Meddows TR, Savory AP, Lloyd RG. RecG helicase promotes DNA double-strand break repair. *Mol Microbiol* 2004;52:119–32.
- [30] Lloyd RG, Rudolph CJ. 25 years on and no end in sight: a perspective on the role of RecG protein. *Curr Genet* 2016. <http://dx.doi.org/10.1007/s00294-016-0589-z>.
- [31] Azeroglu B, Mawer JSP, Cockram CA, White MA, Hasan AMM, Filatenkova M, et al. RecG directs DNA synthesis during double-strand break repair. *PLoS Genet* 2016;12:e1005799.
- [32] Rocha EPC, Comet E, Michel B. Comparative and evolutionary analysis of the bacterial homologous recombination systems. *PLoS Genet* 2005;1:e15.
- [33] Sanchez H, Carrasco B, Cozar MC, Alonso JC. *Bacillus subtilis* RecG branch migration translocase is required for DNA repair and chromosomal segregation. *Mol Microbiol* 2007;65:920–35.
- [34] Bentley SD, Chater KF, Cerdeno-Tarraga A-M, Challis GL, Thomson NR, James KD, et al. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* 2002;417:141–7.
- [35] Ikeda H, Ishikawa J, Hanamoto A, Shinose M, Kikuchi H, Shiba T, et al. Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. *Nat Biotechnol* 2003;21:526–31.
- [36] Huang C-H, Tsai H-H, Tsay Y-G, Chien Y-N, Wang S-L, Cheng M-Y, et al. The telomere system of the *Streptomyces* linear plasmid SCP1 represents a novel class. *Mol Microbiol* 2007;63:1710–8.
- [37] Fischer G, Wenner T, Decaris B, Leblond P. Chromosomal arm replacement generates a high level of intraspecific polymorphism in the terminal inverted repeats of the linear chromosomal DNA of *Streptomyces ambofaciens*. *Proc Natl Acad Sci U S A* 1998;95:14296–301.
- [38] Uchida T, Miyawaki M, Kinashi H. Chromosomal arm replacement in *Streptomyces griseus*. *J Bacteriol* 2003;185:1120–4.
- [39] Volff J-N, Altenbuchner J. Genetic instability of the *Streptomyces* chromosome. *Mol Microbiol* 1998;27:239–46.
- [40] Thibessard A, Leblond P. Subtelomere plasticity in the bacterium *Streptomyces*. In: Subtelomeres. Springer; 2014. p. 243–58.
- [41] Leblond P, Demuyter P, Moutier L, Laakel M, Decaris B, Simonet J-M. Hypervariability, a new phenomenon of genetic instability, related to DNA amplification in *Streptomyces ambofaciens*. *J Bacteriol* 1989;171:419–23.
- [42] Volff J-N, Viell P, Altenbuchner J. Artificial circularization of the chromosome with concomitant deletion of its terminal inverted repeats enhances genetic instability and genome rearrangement in *Streptomyces lividans*. *Mol Gen Genet* 1997;253:753–60.
- [43] Zhang L, Nguyen HC, Chipot L, Piotrowski E, Bertrand C, Thibessard A, et al. The *adnAB* locus, encoding a putative helicase-nuclease activity, is essential in *Streptomyces*. *J Bacteriol* 2014;196:2701–8.
- [44] Huang T-W, Chen CW. A *recA* null mutation may be generated in *Streptomyces coelicolor*. *J Bacteriol* 2006;188:6771–9.
- [45] Pinnert-Sindico S. Une nouvelle espèce de *Streptomyces* productrice d'antibiotiques: *Streptomyces ambofaciens* n. sp. caractères cultureaux. *Ann Inst Pasteur* 1954;87:702–7.
- [46] Hanahan D. Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* 1983;166:557–80.
- [47] MacNeil DJ, Gewain KM, Ruby CL, Dezeny G, Gibbons PH, MacNeil T. Analysis of *Streptomyces avermitilis* genes required for avermectin biosynthesis utilizing a novel integration vector. *Gene* 1992;111:61–8.
- [48] Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* 2000;97:6640–5.
- [49] Thibessard A, Bertrand C, Hiblot J, Piotrowski E, Leblond P. Construction of pDYN6902, a new *Streptomyces* integrative expression vector designed for cloning sequences interfering with *Escherichia coli* viability. *Plasmid* 2015;82:43–9.
- [50] Choulet F, Aigle B, Gallois A, Mangenot S, Gerbaud C, Truong C, et al. Evolution of the terminal regions of the *Streptomyces* linear chromosome. *Mol Biol Evol* 2006;23:2361–9.
- [51] Gust B, Challis GL, Fowler K, Kieser T, Chater KF. PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proc Natl Acad Sci U S A* 2003;100:1541–6.
- [52] Raynal A, Karray F, Tuphile K, Darbon-Rongere E, Pernodet J-L. Excisable cassettes: new tools for functional analysis of *Streptomyces* genomes. *Appl Environ Microbiol* 2006;72:4839–44.
- [53] Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA. *Practical Streptomyces genetics*. Norwich, United Kingdom: John Innes Foundation; 2000.
- [54] Volff JN, Vandewiele D, Simonet JM, Decaris B. Ultraviolet light, mitomycin C and nitrous acid induce genetic instability in *Streptomyces ambofaciens* ATCC23877. *Mutat Res* 1993;287:141–56.
- [55] Bonferroni CE. Il calcolo delle assicurazioni su gruppi di teste. In: *Studi in Onore del Profesor S. O. Carboni*; 1935. p. 13–60.
- [56] Schössler A, Jantos J, Hackmann K, Schrepff H. Characterization of the binding protein-dependent cellobiose and cellobiose transport system of the cellulose degrader *Streptomyces reticuli*. *Appl Environ Microbiol* 1999;65:2636–43.
- [57] Luria SE, Delbrück M. Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* 1943;28:491–511.
- [58] Whitby MC, Ryder L, Lloyd RG. Reverse branch migration of Holliday junctions by RecG protein: a new mechanism for resolution of intermediates in recombination and DNA repair. *Cell* 1993;75:341–50.
- [59] Müller B, West SC. Processing of Holliday junctions by the *Escherichia coli* RuvA, RuvB, RuvC and RecG proteins. *Experientia* 1994;50:216–22.
- [60] Matic I, Rayssiguier C, Radman M. Interspecies gene exchange in bacteria: the role of SOS and mismatch repair systems in evolution of species. *Cell* 1995;80:507–15.
- [61] Magner DB, Blankschien MD, Lee JA, Pennington JM, Lupski JR, Rosenberg SM. RecQ promotes toxic recombination in cells lacking recombination intermediate-removal proteins. *Mol Cell* 2007;26:273–86.
- [62] Carter AS, Tahmaseb K, Compton SA, Matson SW. Resolving Holliday junctions with *Escherichia coli* UvrD helicase. *J Biol Chem* 2012;287:8126–34.
- [63] Lin Z, Kong H, Nei M, Ma H. Origins and evolution of the *recA/RAD51* gene family: evidence for ancient gene duplication and

- endosymbiotic gene transfer. Proc Natl Acad Sci U S A 2006;103: 10328–33.
- [64] Beam CE, Saveson CI, Lovett ST. Role for *radA/sms* in recombination intermediate processing in *Escherichia coli*. J Bacteriol 2002;184: 6836–44.
- [65] Cooper DL, Boyle DC, Lovett ST. Genetic analysis of *Escherichia coli* RadA: functional motifs and genetic interactions. Mol Microbiol 2015; 95:769–79.
- [66] Aravind L, Makarova KS, Koonin EV. Holliday junction resolvases and related nucleases: identification of new families, phyletic distribution and evolutionary trajectories. Nucleic Acids Res 2000;28:3417–32.
- [67] Nautiyal A, Rani PS, Sharples GJ, Muniyappa K. *Mycobacterium tuberculosis* RuvX is a Holliday junction resolvase formed by dimerisation of the monomeric YqgF nuclease domain. Mol Microbiol 2016;100(4): 656–74.

II. Caractérisation des gènes NHEJ-like chez les *Streptomyces*

Multiple and variable NHEJ-like genes are involved in resistance to DNA damage in *Streptomyces ambofaciens*

Hoff G., Bertrand C., Zhang L., Piotrowski E., Chipot L., Bontemps C., Confalonieri F., McGovern S., Lecointe F., Thibessard A., and Leblond P.

Article N°2 : Soumis au journal *Frontiers in Microbiology* le 07/10/16

Résumé des résultats de l'article :

Dans la partie introductive de ce manuscrit de thèse, une synthèse descriptive du mécanisme de NHEJ bactérien et des acteurs qui le composent a été réalisée en se basant majoritairement sur des données obtenues chez la bactérie modèle des actinomycètes, à savoir *Mycobacterium*. Ainsi, il a été constaté que l'intervention du mécanisme de NHEJ est pilotée par la protéine Ku, qui reconnaît la DSB et qui recrute la protéine LigD possédant à la fois une activité nucléase, une activité polymérase et une activité ligase qui interviennent pour le traitement et la ligation des extrémités de la cassure (Pour revue : Bowater et Doherty, 2006 ; Pitcher *et al.*, 2007b ; Shuman et Glickmann, 2007 ; Brissett et Doherty, 2009). En absence de Lig-Dom, la protéine LigC peut toutefois procurer un effet « *back-up* » (Della *et al.*, 2004 ; Bhattarai *et al.*, 2014). Chez les *Streptomyces*, l'identification et la caractérisation de gènes NHEJ-like au sein des génomes complets disponibles sur la base de données NCBI a été initiée avant mon arrivée au sein du laboratoire et poursuivie durant ma thèse.

Pour cela les séquences protéiques Rv0937c et Rv0938 correspondant respectivement à Ku et LigD chez *Mycobacterium tuberculosis* ont été utilisées comme appâts pour les recherches d'identité chez *S. ambofaciens*. De cette manière, nous avons pu constater que la protéine LigD ne possède pas d'homologue en tant que tel dans l'ensemble des génomes étudiés, mais que chaque protéine candidate s'aligne uniquement avec l'un des domaines de LigD. Ainsi, 2 gènes codant un Lig-Dom-like (appelés *ligC* et *ligD*) et 3 gènes codant un Pol-Dom-like (appelés *polK*, *polO*, et *polR*) ont pu être identifiés. En revanche, aucun gène codant pour un Nuc-Dom-like n'a pu être détecté. Il est à noter que LigC_{Sam} apparaît comme étant plus proche de la protéine LigC_{Mtub} que de Lig-Dom_{Mtub}. Trois gènes *ku-like* que nous avons nommés respectivement *kuA*, *kuB* et *kuC* ont également été mis en

évidence. Par la suite, les séquences protéiques qui correspondent aux différents gènes NHEJ-*like* identifiés chez *S. ambofaciens* ont été utilisées comme appâts pour déterminer s'ils sont conservés dans les 38 génomes de *Streptomyces* séquencés. A la suite de ces analyses, les gènes NHEJ-*like* ont pu être répartis dans deux catégories, à savoir une catégorie de gènes « conservés » et une catégorie de gènes « variables ». Le pool de gènes conservés comprend les gènes *kuA*, *ligC*, *polK* et *polR* qui co-localisent systématiquement 2 à 2, formant les loci *kuA-polK* et *ligC-polR*, présents dans l'ensemble des génomes (à l'exception de *S. vietnamensis* dépourvu du locus *kuA-polK*). En revanche, les gènes du pool « variable », à savoir, *kuB*, *kuC*, *polO* et *ligD* sont présents uniquement dans certains génomes.

Par la suite, une collection de mutants délétés pour les gènes NHEJ-*like* a été construite, comprenant notamment des simples mutants ΔkuA , ΔkuB , ΔkuC , $\Delta ligC$, $\Delta ligD$, $\Delta polK$, et $\Delta polR$, mais également des mutants multiples $\Delta kuABC$, $\Delta ligCD$ et $\Delta polKR$. Si l'ensemble des souches mutantes présentent un phénotype colonial similaire à celui de la souche sauvage, leur exposition à différentes doses de rayons électroniques montre en revanche une déficience dans la réponse aux dommages de l'ADN. Cette sensibilité est par ailleurs accrue pour les mutants multiples suggérant une absence de redondance fonctionnelle entre les gènes d'une même famille. Des analyses biochimiques révèlent que les protéines Ku-*like* peuvent protéger les extrémités d'ADN et stimuler leur ligation. Les expériences de RT-qPCR des différents gènes NHEJ-*like* ont révélé une faible expression globale en conditions de laboratoire : si l'expression des gènes *lig* et *pol* s'est avérée trop faible pour être détectée, le taux de transcription des *ku* a en revanche révélé (i) que *kuB* et *kuC* sont faiblement exprimés mais de manière constante le long du cycle de développement tandis que (ii) l'expression de *kuA* augmente dans les phases tardives de croissance. Des analyses de fusion GFP ont également permis de montrer une localisation préférentielle des Ku dans les spores.

Ainsi, les résultats obtenus suggèrent l'existence d'un NHEJ original chez les *Streptomyces*, qui pourrait notamment être actif durant la phase de sporulation. L'implication des gènes NHEJ-*like* « conservés » et « variables » dans la réponse aux dommages de l'ADN soulève de nombreuses questions sur l'existence de plusieurs mécanismes de NHEJ ou sur l'intervention de gènes NHEJ-*like* dans d'autres mécanismes de réparation de l'ADN. L'ensemble de ces éléments est discuté dans l'article ainsi que dans la partie « discussion générale » de ce manuscrit.

Multiple and variable NHEJ-like genes are involved in resistance to DNA damage in *Streptomyces ambofaciens*

Grégory Hoff^{1,2}, Claire Bertrand^{1,2}, Lingli Zhang^{1,2}, Emilie Piotrowski^{1,2}, Ludovic Chipot^{1,2}, Cyril Bontemps^{1,2}, Fabrice Confalonieri⁴, Stephen McGovern³, François Lecointe³, Annabelle Thibessard^{1,2} and Pierre Leblond^{1,2†}

¹ Université de Lorraine, Dynamique des Génomes et Adaptation Microbienne, UMR 1128, Vandœuvre-lès-Nancy, F-54506, France

² INRA, Dynamique des Génomes et Adaptation Microbienne, UMR 1128, Vandœuvre-lès-Nancy, F-54506, France

³ Institut Micalis, INRA, AgroParisTech, Université Paris-Saclay, 78350 Jouy-en-Josas, France.

⁴ Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Université Paris-Sud, Orsay, France.

†corresponding authors. Email: pierre.leblond@univ-lorraine.fr. Phone: 33(0)383684207. Fax: 33(0)383684499.

Running title: NHEJ-like genes in *S. ambofaciens*

ABSTRACT

Non homologous end-joining (NHEJ) is a double strand break (DSB) repair pathway which does not require any homologous template and can ligate two DNA ends together. The basic bacterial NHEJ machinery involves two partners: the Ku protein, a DNA end binding protein for DSB recognition and the multifunctional LigD protein composed a ligase, a nuclease and a polymerase domain, for end processing and ligation of the broken ends. *In silico* analyses performed in the 38 sequenced genomes of *Streptomyces* species revealed the existence of a large panel of NHEJ-like genes. Indeed, *ku* genes or *ligD* domain homologues are scattered throughout the genome in multiple copies and can be distinguished in two categories: the “core” NHEJ gene set constituted of conserved loci and the “variable” NHEJ gene set constituted of NHEJ-like genes present in only a part of the species. Strikingly, deletion of “core” as well as “variable” genes of *Streptomyces ambofaciens* ATCC 23877 led to an increased sensitivity to DNA damage induced by electron beam irradiation. Multiple mutants of *ku*, ligase or polymerase encoding genes showed an aggravated phenotype compared to single mutants suggesting the absence of functional redundancy between genes of the respective families. Biochemical assays revealed the ability of Ku-like proteins to protect and to stimulate ligation of DNA ends. RT-qPCR and GFP fusion experiments suggested that *ku*-like genes show a growth phase dependent expression profile consistent with their involvement in DNA repair during spores formation and/or germination. Finally, our results suggest the existence of a NHEJ pathway in *S. ambofaciens* and its important contribution to DNA damage resistance.

INTRODUCTION

Double strand breaks (DSB) are the most detrimental genomic lesions and their repair constitutes a major challenge for cell viability. In eukaryotes and prokaryotes, two main repair machineries have evolved to repair these DSBs, namely homologous recombination (HR) and Non-homologous end joining (NHEJ) (Deriano and Roth, 2013; Alonso *et al.*, 2013; Glickman, 2014). In contrast to HR which needs an intact template to ensure repair of the break, NHEJ can bind DNA ends together without any template after a facultative DNA ends remodeling step. In multicellular eukaryotes, the NHEJ mechanism represents the predominant DSB repair pathway and requires the intervention of the end tethering heterodimer Ku70/80 as well as diverse end processing proteins (for review: Deriano and Roth, 2013). In prokaryotes, NHEJ is not ubiquitous and the minimal NHEJ machinery first described in *Mycobacterium* consists of two actors: a Ku homologue and an ATP dependent ligase LigD (Weller *et al.*, 2002; Gong *et al.*, 2005; Della *et al.*, 2004). The latter one is a multifunctional protein that carries three distinct enzymatic domains conferring nuclease, polymerase and ligase activities (Pitcher *et al.*, 2005). Unlike DSB repair through HR, DSB repair carried out by NHEJ is mutagenic. Nucleotide addition and/or deletion are scars frequently left after DNA ends ligation (Stephanou *et al.*, 2007; Aniuoku *et al.*, 2008; Gupta *et al.*, 2011). The initial model involving two unique actors is now renewed by the discovery of alternative actors. When the ligase domain of LigD (Lig-Dom) is defective, a back-up ligation activity is provided by LigC, another ATP dependent ligase (Della *et al.*, 2004; Bhattarai *et al.*, 2014). In the same way, *Mycobacterium smegmatis* possesses two additional polymerases which share the same *in vitro* activity than the polymerase domain of LigD (Pol-Dom) (Zhu *et al.*, 2012).

Streptomyces are filamentous bacteria mainly found in soils, sediments and seawater as well as symbionts with plants, fungi and animals (Seipke *et al.*, 2011). Soil is a complex and changing environment where *Streptomyces* have to adapt to the heterogeneousness of the niche (moisture, nutrients, temperature...) and to cope with biotic competition. Their rich and varied secondary metabolism such as signaling molecules or chemical weapons is assumed to mediate their relationships with other organisms (Davies and Davies, 2010; Abrudan *et al.*, 2015). Under limited nutritive conditions, their hyphae differentiate into long chains of exospores providing an efficient way for resistance and dissemination until suitable conditions for another growth cycle are met (Chater, 1993). While hyphae are large *multi-genomic* compartments, spores are meant to include a single genome. In contrast to the vast majority of

bacteria, *Streptomyces* possess a linear chromosome whose size is ranging from 5.96 Mb to 11.94 Mb (<http://www.ncbi.nlm.nih.gov/assembly/?term=Streptomyces>) with a central replication origin and terminally inverted repeats that are covalently bound to specific telomere proteins (Lin *et al.*, 1993). The *Streptomyces* genome presents a high plasticity and their chromosome is highly compartmented, with a conserved central region including mostly housekeeping genes whereas the dispensable genes are predominantly located in the chromosomal arms (Bentley *et al.*, 2002; Ikeda *et al.*, 2003; Kim *et al.*, 2015). The loss of synteny at the ends of the chromosome results from horizontally acquired DNA insertion or deletion events (Choulet *et al.*, 2006; Thibessard and Leblond, 2014). This terminal variability reveals the strong recombination activity taking place at the ends of the genome which is reminiscent of the DNA rearrangements reported several decades ago (Leblond *et al.*, 1990; Volff and Altenbuchner, 1998). Hence, *Streptomyces* are also subject to a high level of genetic instability which is directly linked to major genome rearrangements. Characterization of some of these rearrangements revealed large deletions, chromosome circularization, (Fischer *et al.*, 1997; Inoue *et al.*, 2003), arm replacement (Fischer *et al.*, 1998; Uchida *et al.*, 2003), chromosome end-to-end fusion (Wenner *et al.*, 2003) or large DNA amplifications (Yanai *et al.*, 2006). Molecular analyses showed that chromosome circularization and end-to-end fusion events especially rely on illegitimate recombination whereas chromosome arm replacement or DNA amplifications is dependent on HR. This terminal recombination pattern may result from higher recombination frequencies triggered by the preferential formation of double strand breaks in these regions. On the other hand, genome rearrangements could also be more tolerated if formed in these dispensable regions.

Nevertheless, little is known about DSB repair in *Streptomyces*. Like the other bacteria, *Streptomyces* possess a RecA homologue involved in HR (Aigle *et al.*, 1997; Muth *et al.*, 1997; Vierling *et al.*, 2001; Huang and Chen, 2006). However, recent studies that we performed suggest that *Streptomyces* has an atypical HR pathway. In fact, in contrast to bacterial models like *Escherichia coli*, *Bacillus subtilis* or *Mycobacterium*, deletion of genes encoding a potential helicase-nuclease complex supposed to be involved in the initial end resection step during HR is not viable in *S. ambofaciens* (Zhang *et al.*, 2014). On the other hand, deletion of *ruvABC* and *recG*, two loci encoding factors that have a redundant role in migration and resolution of Holliday junction during late steps of HR, has only a mild effect on recombination efficiency (Hoff *et al.*, 2016). Early bioinformatics studies showed that the two model species, *S. coelicolor* and *S. avermitilis*, both possess a *ku* orthologue suggesting also the existence of a NHEJ pathway in *Streptomyces* (Rocha *et al.*, 2005). Here, we

performed *in silico* analyses in the fully sequenced *Streptomyces* genomes leading to the identification of a large panel of conserved and variable genes suspected to be involved in a NHEJ pathway. Construction of single and multiple NHEJ-like genes mutant strains in *S. ambofaciens* allowed us to study the impact of these different genes in DNA damage repair. All the results we obtained suggest the existence of a complex and atypical bacterial NHEJ pathway in *Streptomyces*.

MATERIAL AND METHODS

Phylogenetic analyses

A total of 38 sequenced *Streptomyces* genomes, including that of *S. ambofaciens* ATCC23877 was scanned to identify genes putatively encoding NHEJ Ku-like or LigD-like proteins based on BLAST similarity search against the nonredundant database. A preliminary BLASTP search (e-value cut-off 10^{-30} ; identity threshold 30%) using *M. tuberculosis* Rv0937c and Rv0938 (Ku and LigD respectively) protein sequences as a query allow to identify and numerate the candidates. Secondly, BLASTP were performed using *S. ambofaciens* ATCC23877 NHEJ-like genes and reciprocal best hits allowed to organize them into groups of orthologues.

For the phylogeny of Ku-like proteins, the tree was built using a maximum-likelihood method based on a GTR + G + I model. Phylogenetic trees were built and edited with Mega6 (Tamura *et al.*, 2013). Support of the tree branches was estimated with 100 bootstrap replicates and all positions with < 80 % site coverage were eliminated. A total of 295 and 309 positions were respectively used for the Ku-like and the LigC/LigD protein phylogenetic trees.

Bacterial strains and plasmids

All strains used in this study are listed in Table 1. The DH5 α *E. coli* strain is used as conservation host for BACs and plasmids. For biochemical analyses, plasmid constructions were done in *E. coli* MC1061 (Casadaban *et al.*, 1980). ER2566 (NEB) was used for *kuB* gene expression and Rosetta (DE3) PLysS (Stratagene) for *kuA* and *kuC* gene expression. *E. coli* strains were grown in Luria-Bertani (LB) medium at 37°C except for the BW25113 pKD20 thermosensitive strain used for PCR targeting, which was grown at 30°C. *S. ambofaciens* strains were grown at 30°C on Soya flour mannitol (SFM) plates except for phenotypical analyses and for liquid cultures which were performed on solid Hickey Tresner (HT) medium. All *Streptomyces* mutants derive from our reference strain *S. ambofaciens* ATCC 23877 (Pinnert-Sindico, 1954). Mutant strains were constructed by PCR targeting as described by Gust *et al.*, 2003. To summarize, a recombinant BAC containing the locus of interest was transformed in the highly recombinogenic *E. coli* BW25113/pKD20 strain in order to replace the target gene by an *oriT*-*aac(3)IV* resistance disruption cassette (Raynal *et al.*, 2006). The ET12567 non methylating strain containing the mobilizing pUZ8002 plasmid was used as donor for intergeneric conjugation of the modified BAC between *E. coli* and *S. ambofaciens* Double-crossover (CO) events leading to the replacement of the target gene by the cassette

were selected by the sensitivity to kanamycin (loss of the BAC vector part) along with the resistance to apramycin (insertion of the resistance cassette). The disruption cassette flanked by *attL* and *attR* sites was excised in the newly obtained mutant strains thanks to the introduction of pOSV508 a plasmid allowing the expression of *int* and *xis* genes from pSAM2 (Raynal *et al.*, 2006). Strains expressing translational EGFP fusions were also constructed by *in vivo* recombination. The *egfp* gene with the transcription terminator from phage λ was cloned from plasmid pIJ8660 (Sun *et al.*, 1999) into an *oriT aac(3)IV* disruption cassette (Raynal *et al.*, 2006). This cassette was then used as the source to make *kuA-egfp*, *kuB-egfp* and *kuC-egfp* fusions. PCR was performed using *egfp-oriT-aac(3)IV* as the template with specific oligonucleotides which anneal to the linker sequence at the 5'end of *egfp* and downstream of the 3'end of *aac(3)IV* and introduce 40 bases of homology to the 3'ends of *kuA*, *kuB* or *kuC*, as well as removing the stop codons. PCR products were introduced into *E. coli* BW25113/pKD20 containing the cognate BACs with either *kuA*, *kuB* or *kuC*, and the following steps of the procedure are described above. In these recombinant strains, the fusion genes (*kuA-egfp*, *kuB-egfp* or *kuC-egfp*, respectively) are the only copies of *kuA*, *kuB* or *kuC*.

Electron beam irradiation (EBI)

An amount of 10^9 spores of different *Streptomyces* strains was diluted in sterile water in 96 well microplates and exposed to 0, 0.1, 0.3, 0.5, 0.7 and 1 kGy of radiation doses. Processing was carried out in a Van de Graaf® electron accelerator (VIVIRAD, Handschuheim, France) in the technological resource center Aérial (Illkirch-Graffenstaden, France). Energy delivered by the electron accelerator reach 2.2 MeV with an intensity of $10\mu\text{A}$. Dose rate equals approximately 60 kGy/h. After irradiation, several spore dilutions were plated with the easySpiral® apparatus (Interscience, Saint-Nom-la-Bretèche, France) on SFM Petri dishes and the number of colonies was counted after 3 days incubation at 30°C. For each genetic context, at least two independent mutant strains were exposed in two independent experiences.

Fluorescent microscopy

Strains for microscopic observations were inoculated in the acute-angle junction of standard-sized microscope coverslips inserted at 45° in SFM agar, and grown for 48 hours. The coverslips were then carefully pulled from the medium and directly mounted in a drop of sterile water on a microscope slide. Fluorescent microscopy observations were carried out using a Axiovert 200M MOT device (Carl Zeiss Micro imaging) equipped with a mercury

vapor lamp and a FITC filter set (470/40 excitation, 540/50 emission). Both contrast phase and fluorescent observations were made with an X100 objective.

DNA preparation

To verify the mutant strains, *Streptomyces* genomic DNA was extracted by low binding sulfate salt method as described by Kieser *et al.*, 2000. Liquid HT medium was used for growth of mycelium before DNA extraction. For biochemical analyses, the gene coding for KuA was amplified from the BAC 16ZB5. *KuB* and *KuC* genes were amplified from the BAC 27ZF1. Primers used for amplification were designed to flank the open reading frames with the coding sequence of a His₆-tag followed by a *NdeI* site on their 5' side, and by a *XhoI* site on their 3' side. The resulting ORFs were cloned under the T7 promoter of the pJ411 vector (DNA2.0) giving the plasmids pSMG258, 259, 260 allowing expression of the proteins 6His-KuA, 6His-KuB, 6His-KuC, respectively (named hereafter KuA, KuB and KuC). Sequences of the primers used in this study are available upon request.

RNA extraction and RT-qPCR analyses

For RNA extraction from cultures on solid medium, streaks were harvested after 48 hours growth at 30°C on HT solid medium covered with cellophanes membranes. For RNA extraction after a genotoxic stress, pre-germinated spores of *Streptomyces* were grown until 0.2 OD600 was reached and exposed to mitomycin C (MMC, 1.5 µg.mL⁻¹, i.e. CMI/4) for 30 min. RNA isolation was performed using the Aurum Total RNA mini-kit (Bio-Rad) according to the manufacturer's instructions. To optimize cell lysis, an additional sonication step (3 x 30 sec) using a Bioruptor apparatus was added. The RNA quality was assessed after migration on 1% agarose gel. In order to eliminate genomic DNA contamination, the RNA samples were treated before reverse transcription with DnaseI (Ambion) in the presence of 10U of RNase inhibitor per µg of RNA. DnaseI was then inactivated at 75°C during 10 minutes in the presence of EDTA. Reverse transcription was performed on 2 µg RNA with an iScript Advanced cDNA synthesis kit (Bio-Rad) according to the manufacturer's instructions. The absence of contamination with genomic DNA was validated by a 30 cycle-PCR on a control aliquot removed before reverse transcription using hrdB-F and hrdB-R primers. Quantitative PCR was undertaken in a CFX96 (Bio-Rad) apparatus in a 10 µL reaction mixture containing respectively 5 µL of SYBR supermix (Bio-Rad), 0.2 µM of each primer and 4 µL of a ten-fold dilution of cDNA. The primers sequences of the tested genes are listed in table 2. The reaction conditions were as follow: a 3 min initial denaturation step at 95°C and 40 cycles of

10 s at 95°C and 30 s at 60°C. In order to check the absence of secondary products, melting curves were realized from 65 to 95°C with an increase rate of 0.5°C/5s. The transcription level of interest genes was analyzed using the mathematical model proposed by Pfaffl (Pfaffl, 2001). The *hrdB* gene, encoding the principal sigma factor in *Streptomyces*, was used as the reference gene for the normalization of the expression level as it was the case in former studies (Galet *et al.*, 2014; Martinez-Burgo *et al.*, 2015).

Proteins purification

Purification of KuA

E. coli Rosetta PlysS cells transformed with pSMG258 were grown at 30°C in 2 liters of LB medium supplemented with 25 mg/mL thymine, 30 µg/mL kanamycin and 25 µg/mL chloramphenicol to OD600 = 0.7. Production of protein was induced with 0.5 mM IPTG for 15 hours at 20°C. Cells were harvested by centrifugation and the pellet was resuspended in 40 mL lysis buffer A (50 mM Tris-HCl pH 8, 0.25 M NaCl). Cells were broken by a freeze / thaw step in liquid N₂ / 37°C and then centrifuged 1 hour at 100 000 g at 4°C. The pellet was resuspended with 40 mL of a 20 mM Tris pH 7.5, 2.5 mM MgCl₂, 0.5 mM CaCl₂, 1 mg/mL DNaseI buffer. After 1 hour at 20°C, the mixture was centrifuged 20 min at 100 000 g at 4°C. All subsequent steps were carried at 4°C. The pellet was washed with 40 ml buffer P₁ (50 mM Tris-HCl, pH 8, 1 M NaCl), and then solubilized with 40 mL of buffer P₂ (50 mM Tris pH 8, 8 M urea). After 1 hour the mixture was centrifuged 20 min at 100 000 g and 3 mL of Ni²⁺ affinity column (Ni-NTA agarose, Qiagen) was added to the supernatant which was transferred onto an Econo-Column (Biorad). The flow through was discarded and the resin was washed with series of 25 mL of buffer P₃ (50 mM Tris pH 8, 1% sucrose, 0.5 M NaCl) with urea concentration decreasing from 8 M to 0 M by steps of 0.5 M. The Ni-NTA column was washed with buffer P₃, then with buffer W (50 mM Tris pH 8.0, 0.5 M NaCl) supplemented with 20 mM imidazole. KuA protein was eluted in buffer W supplemented with 250 mM imidazole and then dialysed against buffer D (50 mM Tris-HCl pH 8, 0.4 M NaCl, 50 % glycerol, 1 mM DTT) prior to storage at -20°C.

Purification of KuC

KuC gene expression was induced in *E. coli* Rosetta PLysS cells transformed with pSMG260 as described for *KuA*. Cells were harvested by centrifugation and the pellet was resuspended in 40 mL lysis buffer B (50 mM Tris-HCl pH 8, 0.25 M NaCl, 0.2 mg/mL lysozyme) supplemented with Complete Antiprotease (Roche) and incubated 1 hour at 20°C. The

mixture was then centrifuged and the pellet was treated with DnaseI, as described for KuA purification. Purification procedure of KuC at this step is the same as the one described for KuA.

Purification of KuB

ER2566 cells transformed with pSMG259 were grown at 30°C in 2 liters of LB medium supplemented with 25 mg/mL thymine and 30 µg/mL kanamycin and *KuB* gene expression was induced as for *KuA* and *KuC* genes. Cells were harvested by centrifugation, the pellet was resuspended in 40 mL of lysis buffer B and incubated 1 hour at 20°C. All subsequent steps were carried at 4°C. Cell lysis mixture was supplemented with 0.5 M NaCl, 0.1% Triton and 0.1 % Nonidet P40 and incubated 1 hour. Two mL of polyethyleneimine (PEI) 10 %, pH 8 was added to precipitate nucleic acids; the mixture was stirred for 1 hour and then centrifuged 1 hour at 100 000 g. The soluble phase was next precipitated by raising the concentration of ammonium sulfate (AS) firstly to 50% of saturation, KuB was mainly found in the supernatant at this step, and secondly to 80% of saturation to precipitate KuB. The corresponding pellet was resuspended in 25 mL of buffer P₁ and 3 mL of Ni-NTA agarose, pre-equilibrated in buffer P₁, were added. The Ni-NTA agarose was washed with buffer P₁, then with buffer W supplemented with 20 mM imidazole and KuB was eluted in buffer W supplemented with 250 mM imidazole. The eluted fraction was dialyzed against buffer D and stored at -20°C.

Final purification step analyzed by SDS PAGE showed that purified KuA, KuB and KuC were obtained at 90% purity and were a mix of full length and truncated forms (Figure S1). Their concentration was determined by UV absorption at $\lambda = 280$ nm on a NanoDrop 1000 Spectrophotometer (Thermo Scientific).

Purification of LigD_{Bsub}

LigD from *Bacillus subtilis* fused in C-terminal to a His₆ tag was purified as previously described (McGovern *et al.*, 2016).

T5 protection assay

Fifty nanograms (~75 fmol) of a 1001 bp linear DNA substrate with 5'-phosphorylated blunt-ends were preincubated with raising concentrations of KuA, KuB or KuC in 19 µL of buffer T5 (50 mM Tris-HCl pH 8, 1 mM DTT, 5 mM MgCl₂, 40 mM NaCl and 5 % glycerol) for 30 minutes at 37°C. One µL of T5 exonuclease (0.5 U/µL, Biolabs) was added to the reaction mixture and reactions were incubated for 30 min at 37°C. Reactions were stopped by adding 4

μL of stop buffer (5 mg/mL proteinase K, 2 % SDS and 0.1 M EDTA) and incubated at 37°C for 30 min. 4 μL of DNA loading buffer (33 % glycerol and 0.25 % xylene cyanol) were added, and the reactions were loaded on a 0.7 % agarose gel in TBE Buffer. After electrophoresis, DNA was stained by the SYBR Gold reagent (Invitrogen) and the gel was photographed using a Chemidoc apparatus (Bio-Rad).

Ligation assay

In a final reaction volume of 20 μL in the T4 DNA ligase ligation buffer (NEB, 50 mM Tris-HCl pH 7.5, 10 mM MgCl_2 , 10 mM DTT, 1 mM ATP), 50 nanograms (~ 75 fmol) of the 1001 bp linear DNA substrate were preincubated with 0.25 or 1 μM of KuA, KuB or KuC for 30 min at 30°C. Then, raising concentrations of LigD_{Bsub} were added to reactions. Ligation reactions were stopped after 2 h at 37°C by adding 4 μL of stop buffer and analyzed as described above.

RESULTS

A “core” and an “variable” NHEJ gene set in *Streptomyces*

Genes homologous to the bacterial NHEJ factors *ku* and *ligD* were sought in the genome of our model strain *Streptomyces ambofaciens* ATCC23877 (Thibessard *et al.*, 2015). To do so, the *M. tuberculosis* Rv0937c and Rv0938 (*Ku_{Mtub}* and *LigD_{Mtub}* respectively) protein sequences were used as query in a BLASTP search. The first observation was that the *LigD_{Mtub}* protein had no full length homologue in *S. ambofaciens*, but that each candidate protein matched only with one single domain at a time. Therefore, the 759 amino acid sequence of *LigD_{Mtub}* was split into 3 slightly overlapping parts (1-300, 271-459 and 451-759) corresponding to the polymerase (Pol-Dom), the nuclease (Nuc-Dom) and the ligase (Lig-Dom) domains, respectively. In this way, 3 Ku-like encoding genes (sharing 36, 39 and 40% of protein identity with *Ku_{Mtub}*), 3 Lig-Dom-like encoding genes (sharing 30, 32 and 41% of protein identity with the Lig-Dom of *LigD_{Mtub}*) and 3 Pol-Dom-like encoding genes (sharing 29, 34 and 41% of protein identity with the Pol-Dom of *LigD_{Mtub}*) were identified in *S. ambofaciens* ATCC23877 genome (Figure 1). No Nuc-Dom-like encoding gene was detected. The 3 *ku*-like genes (*SAM23877_5082*, *SAM23877_6929* and *SAM23877_6942*) were named *kuA*, *kuB* and *kuC* respectively. Two of the 3 Lig-Dom-like genes (*SAM23877_6361* and *SAM23877_0862*) were named *ligC* and *ligD*, respectively. Indeed, *SAM23877_6361* appeared to share more amino-acid identity to *M. tuberculosis* *LigC* (*LigC_{Mtub}*) *i.e.* 60% than to the Lig-Dom of *LigD_{Mtub}* *i.e.* 30% and was therefore considered as *LigC_{Mtub}* orthologue. Similarly, *SAM23877_0862* appeared to be the closest relative of the Lig-Dom of *LigD_{Mtub}*. The third ATP-dependent DNA ligase encoding gene detected (*SAM23877_1283*) would encode a *LigB* protein known to exhibit a strong nick sealing activity *in vitro* and whose defect would not alter mycobacterial NHEJ (Gong *et al.*, 2004). Therefore, this third ligase was not included in our further investigations. The 3 Pol-Dom-like genes (*SAM23877_5081*, *SAM23877_6202* and *SAM23877_6362*) were named *polK*, *polO* and *polR*, respectively. It was noticeable that 2 of the 3 Pol-Dom like genes (*polK* and *polR*) colocalized with *kuA* and *ligC*.

KuA, *KuB* and *KuC* proteins are composed of 365, 383 and 302 amino-acids, respectively. All these proteins harbor the conserved Ku core domain at their N-terminal side, the conserved minimal C-ter domain (red in Figure S2), that has been shown to be required for the interaction with *LigD* in *B. subtilis* (McGovern *et al.*, 2016), and a basic extended C-

terminal domain, which is of 114, 128 and 47 aa, respectively. The pIs of these extended domains (11.64, 10.97 and 10.65) are consistent with those described for the majority of bacterial Ku proteins. This suggests that their roles in DNA binding as well as in the modulation of the ability to thread inward the DNA molecule are conserved in *Streptomyces* (Kushwaha and Grove, 2013; McGovern *et al.*, 2016). Interestingly, a SAP domain, previously predicted in the *S. coelicolor* KuB homologue (*SCO0601*) by Aravind and Koonin (Aravind and Koonin, 2001) is also detected in *S. ambofaciens* KuB (at position 337-370, Figure S2). This SAP domain is assumed to be also involved in nucleic acid binding (Aravind, 2000) and could modify the DNA binding abilities of KuB compared to KuA and KuC. Further experiments are required to highlight the roles and the differences of the extended C-terminal domains of these three Ku proteins.

Once this list of genes potentially involved in a *Streptomyces* NHEJ pathway established, the question of their conservation in the *Streptomyces* genus was addressed by seeking their homologues in 37 other *Streptomyces* complete genome sequences available at <http://www.ncbi.nlm.nih.gov/genome>. Therefore, *S. ambofaciens* KuA, KuB, KuC, LigC, LigD, PolK, PolO and PolR were used as query in BLASTP searches. A wide variety of situations were encountered and a few of them are illustrated in Figure 1. The homologues identified in 38 *Streptomyces* genomes are listed by their locus tag in Table S1 and their presence or absence is summarized in Figure 2. The number and position of NHEJ-like genes varied from one species to another. However, two loci are widely shared among *Streptomyces*: *ligC-polR* which is present in all the investigated genomes and *kuA-polK* which is present in all the tested strains but *Streptomyces vietnamensis*.

Focusing on the main function expected from a functional NHEJ system, we analyzed further the Ku-like and Lig-Dom-like encoding genes and extended our interest to the Pol-Dom-like genes colocalizing with them.

All but *S. vietnamensis* possess at least one *kuA* gene. In contrast, *kuB* was detected in only 8 out of the 38 analysed genomes. Regarding *kuC*, the copy number varies from 0 to 2. Indeed, in the case of *S. coelicolor*, two identical copies of *kuC* are included in the terminal inverted repeats of the SCP1 linear plasmid. A few *Streptomyces* strains carry several copies of *kuA* genes leading us to make a distinction according to their vicinity of *polK*: the neighbours of *polK* were called *kuA* primary genes while standalone *kuA* genes were called secondary *kuA* genes. All the *ku*-like genes (43 *kuA*, 8 *kuB* and 17 *kuC*) were aligned and submitted to a Maximum Likelihood phylogenetic analysis. The tree topology (Figure 3), together with the

fact that the primary *kuA* gene is almost ubiquitous, suggested that KuA was probably present in the *Streptomyces* common ancestor and evolved along with *Streptomyces*. The secondary *kuA* gene carried by *Streptomyces* sp. 769, *S. albulus* ZPM, *S. albulus* NK660 and *S. lydicus* shares an average nucleotidic identity ranging from 67 to 72% with their respective primary copy. For these species, the primary and secondary KuA phylogenies are congruent, suggesting that their common ancestor gained a second *kuA* copy either by gene duplication or by gene transfer. Regarding the *kuC* gene, it was present in 17 out of 38 analyzed genomes, 7 of which were plasmid-borne. It was closely related to KuA and formed a sister group of KuA within a large KuA-KuC monophyletic group that excluded KuB. In fact *kuB* genes formed a strong monophyletic group and appeared much more scattered in *Streptomyces* genera.

Regarding the distribution of LigC and LigD ligases, LigC encoding genes were found in all genomes and always in the close vicinity of *polR* as in *M. tuberculosis* (Figure 1). In contrast, LigD was detected in 29 out of the 38 cases (Figure 2). In 6 of them the encoding gene colocalized with *kuC*, and in *S. cattleya*, it colocalized with *kuA-polK* (see Figure 1). Phylogenetic analysis by Maximum Likelihood method (Figure S3) showed that the LigC proteins branched together with *M. tuberculosis* LigC with a topology that seemed quite congruent with the species relations, suggesting a vertical inheritance of this conserved locus. LigD proteins also branched together and together with Lig-Dom of *M. tuberculosis* LigD.

Altogether, two NHEJ-like gene categories were delineated: (i) those that are conserved among the *Streptomyces* and were considered as “core” putative NHEJ factors (*i.e.* *kuA* and the colocalized *polK*, *ligC* and the colocalized *polR*) and (ii) those that are sporadically represented among the *Streptomyces* and were considered as “variable” putative NHEJ factors (*i.e.* *kuB*, *kuC* and *ligD*).

NHEJ-like genes are not essential in *S. ambofaciens*

After *in silico* identification of the “core” and “variable” NHEJ putative factors in *Streptomyces*, we constructed deleted mutant strains by replacing the chromosomal locus with an excisionable antibiotic resistance cassette (Raynal *et al.*, 2006) by the use of PCR targeting (Gust *et al.*, 2003). The whole selection procedure was performed several times simultaneously in order to select at least two independent mutant lineages for each genetic context. Table 1 summarizes the different genotypes constructed for this study. The mutations were checked by PCR amplification of the mutated locus. The maintenance of the

chromosomal linearity of the mutants was checked by PCR amplification of the telomeres and of some loci scattered along the chromosome (not shown).

All the mutations, simple or multiple, were obtained with classical frequencies indicating that none of the targeted genes (“core” or “variable”) was essential. In addition, there was no significant difference between the wild-type reference strain and any of mutants in colonial morphology and sporulation under normal growth conditions (not shown).

Implication of NHEJ-like genes in DNA damage response

Electronic beam irradiation (EBI) was shown to induce loss of viability of *Bacillus* spores by damaging spore coat, by altering membrane permeability or by DNA fragmentation (Fiester *et al.*, 2012). High level energy beam such as those used in this work (2.2 MeV) is likely to trigger double strand breaks (DSB) as the most significant lesions in the spore embedded DNA. Efficient germination will require DNA repair in the absence of an intact copy of the broken DNA sequence, and thus would involve NHEJ. In order to test the involvement of the NHEJ-like factors to DNA damage response, spores of the wild-type *S. ambofaciens* reference strain as well as spores of at least two independent mutants for each genetic context were irradiated by EBI from 0.1 to 1 kGy and their survival was assessed after growth on SFM medium. The results are represented in Figure 4.

Both “core” and “variable” ku genes are involved in DNA damage response

Depending on the dose, deletion of *kuA*, which belongs to the “core” NHEJ-like gene set, conferred a 2 to 4 fold increase of the sensitivity compared to the wild-type (Figure 4A). Since *kuA* is present in the large majority of the sequenced genomes, it appears as the major Ku actor of the putative *Streptomyces* NHEJ. Surprisingly, mutations of the variable *ku*-like genes, *kuB* and *kuC*, also conferred an increased sensitivity but at a lower level, with a maximal 2 fold cell viability decrease reached after 0.5 and 0.7 kGy exposition for both contexts (Figure 4A). In addition, the sensitivity was aggravated when *kuB* and *kuC* were both deleted in a ΔkuA background to generate a triple $\Delta kuABC$ mutant, until reaching a 7 fold cell viability diminution in comparison to the WT for the medium radiation doses (Figure 4A). This result suggests that the three Ku products participate to the response to DNA damaging agents in spores, each of them assuming their share of the DNA repair independently from the two others. Additionally, the triple $\Delta kuABC$ mutation assumed to abolish any putative NHEJ repair strongly increased the sensitivity of a HR deficient context (*i.e.* $\Delta recA$) (Figure 4D).

This last result shows that RecA and Ku-like proteins are both important contributors to the *S. ambofaciens* spore resistance to EBI.

Two ATP dependent ligases involved in DNA damage response

In *Mycobacterium*, LigD is the main NHEJ ligase whereas LigC only provides a back-up activity (Della *et al.*, 2004). In contrast to *Mycobacterium* where a single *ligC* mutation does not impact survival rate to DNA damage, the defect of both the conserved *ligC* and the variable *ligD* of *S. ambofaciens* conferred a significant sensitivity to EBI. Surprisingly, the deletion of *ligD* conferred an even higher sensitivity than the mutation of *ligC* (Figure 4B). Furthermore, the $\Delta ligCD$ double mutant showed an aggravated phenotype in the same proportions as a $\Delta kuABC$ triple mutant (Figure 4B). These data clearly show that the two ligase activities participate to DNA repair in response to EBI.

Pol-Dom genes are involved in DNA damage response

Single deletion of *polK* or *polR*, the two conserved homologues of *Mycobacterium* Pol-Dom colocalized respectively with *kuA* and *ligC* led to a weak sensitivity to electronic beam exposure. For both single mutants, increase of sensitivity in comparison to the wild type does not exceed 1.5 fold for each dose (Figure 4C). On the other hand, double *polRK* mutant strains are approximately twice more sensitive than the single mutants (Figure 4C). Both loci are therefore involved in DNA damage response.

The *ku* genes are differentially expressed during development

In order to investigate the spatial and temporal expression of NHEJ-like genes during development, a time-course of RT-PCR experiments was performed and translational fusions with EGFP were observed by fluorescent microscopy. First, RNAs were extracted from WT growing strains 24 h, 36 h, 48 h and 60 h after streaking spores on HT solid medium. Reverse transcription was performed and the RNA levels of NHEJ-like genes were assessed by quantitative PCR. The expression level of the targeted genes was normalized to *Streptomyces* reference *hrdB* gene (encoding the main sigma factor in *Streptomyces*). First of all, the results showed that NHEJ-like genes are weakly expressed throughout the development. For instance, the expression of *ligC*, *ligD*, *polK* and *polR* was too low to be monitored by qPCR whereas the expression of *ku* genes was 10 to 50 fold lower than *hrdB* expression (Figure 5A). Secondly, *kuA* expression was clearly growth-phase dependent with a 10 fold increase between 24 h and 60 h after germination. It is important to indicate that sporulation began to

be observable after 48 h of development in these growth conditions. There was no difference throughout the time-course for *kuB* and *kuC* expression. Both of them were less expressed than *kuA* at the late stages of development. The WT strain was also grown to OD600 of 0.2 in HT liquid medium and then exposed to MMC ($1.5 \mu\text{g}\cdot\text{ml}^{-1}$) for 30 min. RNAs were extracted and reverse-transcribed to assess by PCR the regulation of NHEJ-like genes during a genotoxic stress. No induction or repression of NHEJ-like genes was noticeable, while the positive induction control (*recA*) was up-regulated up to 20 fold (not shown).

To observe the cellular localization of Ku proteins in physiological conditions, *egfp* translational fusions were constructed at *kuA*, *kuB* and *kuC* chromosomal loci. Fusion strains were inoculated on SFM agar plates and grown for 48 h before observations by fluorescent microscopy along with a negative control strain without any *egfp* gene. A positive control strain constitutively expressing *egfp* gene exhibited a strong green fluorescence signal in all cells throughout the developmental stages of the bacteria, proving the reporter gene to be functional in *S. ambofaciens* (not shown). For KuA fusion, the green fluorescence signal was significant in spores and segmenting aerial hyphae but not in vegetative mycelium (Figure 5B), showing a specific accumulation of KuA in spores. This result corroborates the phase dependent expression pattern of *kuA* observed in the previous qPCR analysis. For KuB and KuC fusions, although the signals were weaker than for KuA fusion, they also revealed the accumulation of KuB and KuC in spore chains. Since no significant modulation of gene expression was detected for *kuB* and *kuC* along the growth cycle (Figure 5A), this protein accumulation pattern suggests a translational (and/or a post-translational) regulation level.

The Ku proteins protect linear dsDNA from the T5 exonuclease and stimulate the ligase activity of LigD_{Bsub}.

To assess which proteins are required in a putative *S. ambofaciens* NHEJ, we tried to purify all the proteins putatively implicated in this pathway. Among them, we succeeded to express and purified *S. ambofaciens* KuA, KuB and KuC proteins (39.1 KDa, 42.4 KDa and 33.6 KDa respectively). To evaluate the oligomeric state of the three Ku proteins, we performed gel filtration experiments. KuA and KuB were eluted in fractions corresponding to a molecular weight of 160 and 100 kDa, respectively, strongly suggesting that KuA and KuB are multimeric (Figure S1). To date, characterized bacterial Ku are dimers. However the elongated shape of the *B. subtilis* Ku (Ku_{Bsub}) dimer (Mc Govern *et al.*, 2016) led to the elution of this protein with apparent molecular weight corresponding to a tetramer. Then KuA and KuB could also be dimers displaying elongated shape, as for Ku_{Bsub}. KuC was clearly

purified in different oligomeric states ranging from 70 to more than 600 kDa. The main elution peak, corresponding to a protein of apparent molecular weight around 160 kDa, is similar to the one observed for KuA.

KuA, KuB and KuC were able to protect a linear dsDNA molecule from the T5 exonuclease (Figure 6A), a property shared by the Ku_{Bsub} protein (Mc Govern *et al.*, 2016) and the *Pseudomonas aeruginosa* Ku against different exonucleases (Zhu and Shuman, 2010). This result suggests that these proteins are able to bind DNA ends. Another role of Ku in the bacterial NHEJ is to recruit the LigD at DNA ends and to stimulate its ligase activity. Remarkably, KuA, KuB and KuC were able to stimulate the ligation activity of the *B. subtilis* LigD protein (Figure 6B), even if they were less efficient than Ku_{Bsub}. However, we observed a different concentration-dependent stimulation of the ligase by the *S. ambofaciens* Ku proteins. Raising the concentration of KuA and KuC from 0.25 to 1 μ M in the ligation assay led to slightly increase ligation efficiency. At the opposite, the stimulation activity observed with 0.25 μ M of KuC is abolished when its concentration was raised to 1 μ M. A similar inhibitory effect was observed with Ku_{Bsub}.

DISCUSSION

Streptomyces NHEJ-like gene organization is atypical

The presence of a NHEJ pathway in prokaryotes appears to be limited to some phyla representing about 20% of the species (estimated on a sample of 2645 complete bacterial genomes) (Mc Govern *et al.*, 2016). The minimal set of NHEJ gene encodes a two partner system including the homodimeric Ku protein and a ligase polypeptide composed of the ligase enzymatic domain (Lig-Dom) fused to a nuclease (Nuc-Dom) and a polymerase (Pol-Dom) domain. The latter enzymatic activities process DNA ends before ligation. The NHEJ gene repertoire ranges from the simplest in *P. aeruginosa* with a single Ku and a ligase (ligase D) to the more complex with those described in *Agrobacterium* (Zhu and Shuman, 2007) or *Sinorhizobium* (Kobayashi *et al.*, 2008) and *Streptomyces* as described in this work. Hence, while *B. subtilis* and *Mycobacterium* possess a single Ku, a single LigD and one or two LigC, *A. tumefaciens* possesses three Ku, two LigD and three LigC (Zhu and Shuman, 2007). *Sinorhizobium meliloti* and other rhizobia possess up to 4 Ku, 5 LigD and one LigC (Kobayashi *et al.*, 2008); the ligase, nuclease and polymerase domains are present in different organizations, sometimes (i) fused together in a single polypeptide or (ii) encoded by distinct genes. This situation is similar in archaea where the putative ligase, nuclease and polymerase domains are encoded by distinct genes however kept into an operonic organization assumed to ensure the co-regulation of the activities (Bartlett *et al.*, 2013). However, among actinobacteria, *Streptomyces* appears as an exception, since other main phyla (*Nocardia*, *Frankia*, *Rhodococcus* and *Corynebacterium*) seem to possess a multidomain LigD homologue (not shown).

In *Streptomyces*, several levels of complexity are also noticeable. First, the activities putatively involved in NHEJ are encoded by standalone genes that are scattered along the genome. Consequently the repertoire is rather expanded: the number of genes ranging from none up to 4 for *ku*, from one to two Lig-Dom encoding genes (*ligC*, *ligD*), from 2 to 3 Pol-Dom encoding genes, from none to one Nuc-Dom encoding gene. The total number of NHEJ-like genes ranges from 4 in *S. bingchenggensis* and *S. venezuelae* to 8 in *S. ambofaciens* and *S. reticuli*: 3 *ku* (*kuA*, *kuB*, *kuC*), one *ligD* and one *ligC*, and 3 polymerase genes (*polK*, *polR*, *polO*). These genes can fall into two groups. On one hand, the “core” set includes *kuA*, *ligC*, *polK* and *polR* which are present in almost all *Streptomyces* species and organized in two loci of rather conserved organization (i.e. *kuA-polK*, *ligC-polR*). The only exception appears to be in *S. vietnamiensis* which lacks the couple *kuA-polK* and the other *ku* genes. On the other

hand, the rest of the genes constitute the “variable” gene set. An interesting feature to highlight is a gene duplication event involving *kuA* or *kuC*. The duplicated copies are either located in the variable region of the genome or present within the terminal inverted repeats of the linear replicons. This phenomenon is interesting since duplicated copies can evolve through gene divergence to raise new function. This mechanism is rather rare in bacteria but gene duplication seems frequent in *Streptomyces*, notably with large terminal duplication probably resulting from DSB repair by a mechanism related to break-induced replication.

Involvement of the “core” and “variable” NHEJ-like gene sets in DNA repair

The large number of NHEJ-like genes in *Streptomyces* as well as their variability questions their implication in DNA repair mechanisms. It is tempting to imply the “core” NHEJ gene set in a classical NHEJ pathway with a Ku (*i.e.* KuA) and an ATP-dependent DNA ligase (*i.e.* LigC). In addition to being the conserved *ku* gene, *kuA* is the more expressed of the 3 *ku*-like genes in the late growth phases, its defect conferred the most marked sensitivity to EBI and finally, KuA is the most efficient of the 3 Ku-like proteins in stimulating LigD_{Bsub} activity *in vitro*.

It might be surprising that the NHEJ ligase could be LigC rather than LigD which is the canonical NHEJ ligase in other bacterial species. However it should be reminded that in *Streptomyces*, LigD carries the only ligase domain, as does LigC. Both ligases are devoid of the Pol-Dom domain known to be involved in its recruitment by Ku. Therefore LigD and LigC may ensure similar functional roles in *Streptomyces*. One could hypothesize that a gene rearrangement split the ligase domain from the two others (*i.e.* Pol-Dom and Nuc-Dom) in a *Streptomyces* common ancestor. The two ligase genes would have then provided an equivalent contribution to the bacterium, finally resulting in the replacement of LigD by LigC in the “core” NHEJ set by a founder effect.

In *S. ambofaciens*, in which both *ligC* and *ligD* cohabits, single mutants displayed a sensitive phenotype to EBI and the double mutant displayed an aggravated phenotype. In this context, at least two hypotheses can be drawn. First, both LigC and LigD could participate to the same repair pathway; in that case, the sensitivity of single mutants and the aggravated phenotype of double mutants could be explained by a dose effect that is the need for a certain amount of ligases to cope for the induced DNA damage. In that case, both ligase genes would contribute to achieve the required level of ligase. Alternatively, they could participate to different repair pathways required to take over varied DNA damage (see below).

If the “core” NHEJ gene set indeed establishes a functional NHEJ pathway, the presence of two Pol-Doms in the “core” NHEJ gene set constitutes another intriguing feature, especially considering that they do not seem to have redundant functions. Hence, the sensitivity of both *polK* and *polR* simple mutants was aggravated in the double mutant. As shown in *M. tuberculosis* NHEJ, the role of the Pol-Dom domain of LigD would be less to provide a polymerase activity than to recruit LigD to Ku (Pitcher *et al.*, 2005). Then PolK and/or PolR may play a role of linker/adaptor to recruit a ligase; the more likely being that PolK and PolR recruit LigC, however on different types of DNA extremities (*e.g.* blunt, 3’ or 5’ overhanging ends). In the case of the presence of LigD (as in *S. ambofaciens* for example), the two PolK and PolR may recruit both LigC and LigD.

It is particularly noticeable that the “variable” genes (*i.e.* *ligD* but also *kuB*, *kuC*) were involved in the response to Electron Beam Irradiation (EBI). The exposure to EBI is assumed to trigger DSBs. However, different types of lesions can result from direct transmission of energy to the DNA structure and from secondary oxidative stress (*e.g.* oxidized bases). In consequence, damage induced by EBI requires NHEJ but may also require other DNA repair pathways. Thus, these “variable” genes would (i) support the NHEJ encoded by the “core” gene set in an opportunistic way, (ii) constitute an alternative NHEJ pathways or (iii) constitute a repair pathway dedicated to other DNA damage single strand gaps or abasic sites). The first hypothesis would imply a dose effect of the Ku as well as the ligase proteins to explain the aggravated phenotype of the multiple mutants (*i.e.* $\Delta kuABC$ and $\Delta ligCD$). In this first hypothesis, considering that prokaryote Ku act as dimers (McGovern *et al.*, 2016), we can speculate that these three Ku proteins could act as homodimers as well as heterodimers. A data in favour of the second hypothesis (an alternative NHEJ pathway) is the colocalization of the *kuC* and *ligD* genes in 6 *Streptomyces* species clustering with *S. griseus* (Table S1 and Figure 3). However, no clear trend towards the co-occurrence of these two “variable” genes can be noticed in the 38 analysed genomes. In *B. subtilis*, the recent involvement of LigD_{Bsub} and Ku_{Bsub} in base excision repair along with their major role in NHEJ supports the third hypothesis (de Ory *et al.*, 2014; de Ory *et al.*, 2016).

In order to solve the question of the involvement of these genes in a NHEJ pathway, a forward-looking approach could be to specifically induce DSB formation at targeted chromosomal loci and assessed the sensitivity of the NHEJ-candidate genes to identify those truly involved in DSB repair *per se*.

A late growth phase role for NHEJ in Streptomyces

The presence of a NHEJ pathway in bacteria is usually assigned to DSB repair in low rate growth phases characterizing bacterial species involved in pathogenic processes or exposed to harsh environmental living conditions. When bacteria replicate such as in laboratory conditions, sister chromosomes providing an intact template are preferentially used to ensure faithful DSB repair by homologous recombination. However, when replication is slowing down (*i.e.* stationary phases) or stopping (*e.g.* in resistance forms such as spores), DSB repair strongly relies on NHEJ. This is consistent with the fact that NHEJ genes are expressed preferentially when DSB repair cannot be ensured by a faithful process. Hence, in *B. subtilis* as well as in *Streptomyces*, the NHEJ genes were shown to be expressed in stationary phase and to accumulate in spores. The forespore-specific sigma factor σ^G in *B. subtilis* directly regulates *ykoU* (coding for LigD) and *ykoV* (coding for Ku), inducing their expression during the forespore development (Wang *et al.*, 2006). In addition, *B. subtilis* Ku protein specifically localizes to the nucleoid of germinating spores while it disappears at later stages of germination. In *Streptomyces*, gene transcription analyses showed in this work an expression of the *ku* in stationary phase with an increase of *kuA* expression during sporulation. Moreover, GFP fusions showed a predominant localization of the Ku proteins in spores. Furthermore, mutations in *ku* and ligase genes confer sensitivity to genotoxic agents to *B. subtilis* as well as *Streptomyces* spores. Sensitivity to heat and desiccation which are biotic parameters known to trigger DSB is also associated to the NHEJ mutations in *B. subtilis* and *Mycobacterium smegmatis* (Moeller *et al.*, 2007; Pitcher *et al.*, 2007; Stephanou *et al.*, 2007). The presence of multiple *ku* genes (and of other NHEJ-like genes) in *Sinorhizobium meliloti* is assumed to correspond to the need for DNA repair at different steps of the cell cycle. Although no induction of the NHEJ pathways was observed under genotoxic conditions (*i.e.*, IR exposure), the expression patterns of the *sku* genes suggest that each Ku protein has a distinct role under different conditions (free bacterial form, bacteroids). In contrast, NHEJ would not be required during the establishment of symbiosis since a strain lacking all four *sku* genes was shown to establish functional nodules (Kobayashi *et al.*, 2008). As in *S. meliloti*, it was shown that each of the *S. ambifaciens* *ku* genes is functional since their deletion conferred a phenotype of sensitivity towards genotoxic agents. Mutant strains lacking *kuA*, *kuB* and *kuC* genes showed an aggravated phenotype suggesting that each participated to the response to DNA damage. We can therefore suppose that like in *S. meliloti*, intervention of the *Streptomyces* *ku* genes is done under specific conditions.

Recently in *Pseudomonas putida*, the NHEJ was shown to participate to another late growth phase phenomenon (Paris *et al.*, 2015). Hence, bacteria frequently encountered nutrient starvation during late-growth phase. Populations starving display a phenomenon called stationary-phase mutagenesis (also known as adaptive mutations) from which emerge mutants able to overcome the shortage and initiate new vegetative cycles. The underlying mechanisms notably involve unfaithful DNA polymerases during homologous recombination events (Shee *et al.*, 2012). Although the involvement of NHEJ in late mutagenesis remains unknown, deficiency of each of the NHEJ activities triggered an altered pattern of mutations (Paris *et al.*, 2015). The authors speculate on ability to bypass DNA lesions of the LigD polymerase domain.

Origin of the “core” and “variable” NHEJ gene sets in Streptomyces

While the origin of the NHEJ “core” gene set is likely to result from vertical inheritance from a common *Streptomyces* ancestor, the origin of the “variable” set is questionable. The sporadic distribution of the “variable” genes may reflect either the existence of NHEJ in a bacterial common ancestor and its subsequent sporadic loss during genus/species separation. Reciprocally, it may reflect the acquisition of NHEJ genes through horizontal gene transfer in some phyla. Both hypotheses can also be combined and result in the observed complexity. The most explicit argument about their origin is their localization in the variable regions of the chromosome or in plasmids. While plasmids can be mobilized and even self-transmitted, the terminal regions of the *Streptomyces* chromosome is known to be targeted by recombination phenomena leading to insertions of acquired DNA and further losses (Choulet *et al.*, 2006).

Impact of NHEJ on genome evolution

Whatever the evolutionary past of the NHEJ gene sets, the presence of a conserved “core” gene set enriched by additional functional “variable” genes may confer some advantage. In all bacteria studied so far, NHEJ is not essential (Ku/LigD) for growth, at least under laboratory conditions. Mutants of NHEJ in *Mycobacterium*, *Pseudomonas* or *Bacillus* are not affected although they show sensitivity to environmental stress (Gong *et al.*, 2004; Gong *et al.*, 2005; Weller *et al.*, 2002; Jacobs *et al.*, 2003). We came to the same conclusion in *S. ambofaciens* although it was particularly interesting to note that a part of the “variable” genes (*i.e.* *kuB*, *kuC* and *ligD*) was involved in response to genotoxic agents. Hence the aggravated phenotype observed in double or multiple mutant strains (*i.e.* *kuABC*, *ligCD*) revealed the cooperative actions of the multiple alleles. The “variable” genes in *Streptomyces* may increase the

flexibility and/or efficiency of the DSB repair mechanism and influence evolution of the genome. Lessons about the role of NHEJ in genome evolution can be drawn from organisms lacking NHEJ. Hence, the only eukaryote lacking NHEJ is the yeast *Lachancea kluyveri* (Gordon *et al.*, 2011). Compared to other species of the Saccharomycetaceae family, this species showed a significantly low number of genomic rearrangements and the lack of telomere-to-telomere fusions. The authors suggest that the loss of the NHEJ and an alternative end joining pathway (named microhomology-mediated end joining (MMEJ) or alternative end joining (A-EJ)) is responsible for the low recombination rate assessed in this phylum. In this way, we can hypothesize that in *Streptomyces* the presence/absence of NHEJ genes in addition to the “core” NHEJ (the “variable” genes participating to a unique NHEJ or other alternative pathways) may modulate the frequency of DNA rearrangements and DNA acquisition through horizontal gene transfer and confer different evolution rates to *Streptomyces* species. These genes being themselves subject to transfer, the recombinant phenotype (*i.e.* hyperrecombinant) could be variable within a population and favour genetic diversification and adaptation to the environment. This situation could be reminiscent of mutator strains identified in *E. coli* (for review: Giraud *et al.*, 2001) but applied to the capacity to recombine DNA including incoming genetic information.

Future perspectives

A key question is the relative contribution of HR and NHEJ to the evolution and the remarkable compartmentalization of the linear *Streptomyces* chromosome. Is there any regulation between the two pathways during the development of the bacteria? In *S. avermitilis*, the deletion of the *ku* homologues (*ku1* and *ku2* corresponding to *kuA* and *kuC* respectively) was associated to an increased efficiency of homologous recombination (Zhang *et al.*, 2012). This may reveal the existence of a balance resulting from the competition between HR and NHEJ repair mechanisms, and that both repair pathways co-exist at least during replicating phases. Is there any spatial regulation of the DSB repair mechanisms? DSB could be healed by different mechanisms depending of their chromosomal location leading to either faithful repair or possible integration of foreign information.

AUTHOR CONTRIBUTIONS

Investigation: GH, CB, LZ, EP, LC, SM; Computer analysis: AT, CyB; Methodology: CyB, FC, FL; Conceptualization: CB, AT, FL, PL; Writing original draft: GH, CB, AT, PL; Writing-review and editing: FL, FC, CyB, CB, AT, PL; Resources: FC, FL, PL; Funding acquisition: CB, AT, PL; Supervision: AT, PL.

Since there is no middle initial to differentiate Claire Bertrand and Cyril Bontemps, the second one was mentioned as CyB.

FUNDING

This work was supported by the French National Research Agency program Streptoflux (ANR Blanc 0096_01), by ANR through the Laboratory of Excellence ARBRE (ANR-12-LABXARBRE-01) and by the Région Lorraine.

ACKNOWLEDGMENTS

We would like to thank Sébastien Riegler and Florian Kuntz from the Technological Resource center Aérial (Illkirch-Graffenstaden, France) for their technical support and their warm welcome.

BIBLIOGRAPHY

- Abrudan, M. I., Smakman, F., Grimbergen, A. J., Westhoff, S., Miller, E. L., van Wezel, G. P., *et al.* (2015). Socially mediated induction and suppression of antibiosis during bacterial coexistence. *Proc. Natl. Acad. Sci. U. S. A.* 112, 11054–11059. doi:10.1073/pnas.1504076112.
- Aigle, B., Holl, A., Angulo, J. F., Leblond, P., and Decaris, B. (1997). Characterization of two *Streptomyces ambofaciens recA* mutants: identification of the RecA protein by immunoblotting. *FEMS Microbiol. Lett.* 149, 181–187.
- Alonso, J. C., Cardenas, P. P., Sanchez, H., Hejna, J., Suzuki, Y., and Takeyasu, K. (2013). Early steps of double-strand break repair in *Bacillus subtilis*. *DNA Repair* 12, 162–176. doi:10.1016/j.dnarep.2012.12.005.
- Aniukwu, J., Glickman, M. S., and Shuman, S. (2008). The pathways and outcomes of mycobacterial NHEJ depend on the structure of the broken DNA ends. *Genes Dev.* 22, 512–527. doi:10.1101/gad.1631908.
- Aravind, L., and Koonin, E. V. (2001). Prokaryotic homologs of the eukaryotic DNA-end-binding protein Ku, novel domains in the Ku protein and prediction of a prokaryotic double-strand break repair system. *Genome Res.* 11, 1365–1374. doi:10.1101/gr.181001.
- Aravind, L., Makarova, K. S., and Koonin, E. V. (2000). Holliday junction resolvases and related nucleases: identification of new families, phyletic distribution and evolutionary trajectories. *Nucleic Acids Res.* 28(18), 3417–3422.
- Bartlett, E. J., Brissett, N. C., and Doherty, A. J. (2013). Ribonucleolytic resection is required for repair of strand displaced nonhomologous end-joining intermediates. *Proc. Natl. Acad. Sci. U. S. A.* 110, E1984–E1991. doi:10.1073/pnas.1302616110.
- Bentley, S. D., Chater, K. F., Cerdeno-Tarraga, A.-M., Challis, G. L., Thomson, N. R., James, K. D., *et al.* (2002). Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3 (2). *Nature* 417, 141–147.
- Bhattacharai, H., Gupta, R., and Glickman, M. S. (2014). DNA ligase C1 mediates the LigD-independent nonhomologous end-joining pathway of *Mycobacterium smegmatis*. *J. Bacteriol.* 196, 3366–3376. doi:10.1128/JB.01832-14.
- Casadaban, M. J., and Cohen, S. N. (1980). Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. *J. Mol. Biol.* 138(2), 179–207.
- Chater, K. F. (1993). Genetics of differentiation in *Streptomyces*. *Annu. Rev. Microbiol.* 47, 685–711.
- Chayot, R., Montagne, B., Mazel, D., and Ricchetti, M. (2010). An end-joining repair mechanism in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* 107, 2141–2146. doi:10.1073/pnas.0906355107.

- Choulet, F., Aigle, B., Gallois, A., Mangenot, S., Gerbaud, C., Truong, C., *et al.* (2006). Evolution of the terminal regions of the *Streptomyces* linear chromosome. *Mol. Biol. Evol.* 23, 2361–2369. doi:10.1093/molbev/msl108.
- Davies, J., and Davies, D. (2010). Origins and evolution of antibiotic resistance. *Microbiol. Mol. Biol. Rev.* 74, 417–433. doi:10.1128/MMBR.00016-10.
- Della, M., Palmboos P. L., Tseng H., Tonkin L. M., Daley J. M., Topper L. M. *et al.* (2004). Mycobacterial Ku and ligase proteins constitute a two-component NHEJ repair machine. *Science* 306, 683–685. doi:10.1126/science.1099824.
- Deriano, L., and Roth, D. B. (2013). Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage. *Annu. Rev. Genet.* 47, 433–455. doi:10.1146/annurev-genet-110711-155540.
- Fiester, S. E., Helfinstine, S. L., Redfearn, J. C., Uribe, R. M., and Woolverton, C. J. (2012). Electron beam irradiation dose dependently damages the *Bacillus* spore coat and spore membrane. *Int. J. Microbiol.* 2012, 1–9. doi:10.1155/2012/579593.
- Fischer, G., Decaris, B., and Leblond, P. (1997). Occurrence of deletions, associated with genetic instability in *Streptomyces ambofaciens*, is independent of the linearity of the chromosomal DNA. *J. Bacteriol.* 179, 4553–4558.
- Fischer, G., Wenner, T., Decaris, B., and Leblond, P. (1998). Chromosomal arm replacement generates a high level of intraspecific polymorphism in the terminal inverted repeats of the linear chromosomal DNA of *Streptomyces ambofaciens*. *Proc. Natl. Acad. Sci. U. S. A.* 95, 14296–14301.
- Galet, J., Deveau, A., Hôtel, L., Frey-Klett, P., Leblond, P., and Aigle, B. (2015). *Pseudomonas fluorescens* pirates both ferrioxamine and ferricoelichelin siderophores from *Streptomyces ambofaciens*. *Appl. Environ. Microbiol.* 81, 3132–3141. doi:10.1128/AEM.03520-14.
- Giraud, A., Radman, M., Matic, I., and Taddei, F. (2001). The rise and fall of mutator bacteria. *Curr. Opin. Microbiol.* 4, 582–585.
- Glickman, M. S. (2014). Double-strand DNA break repair in mycobacteria. *Microbiol. Spectr.* 2(5):MGM2-0024-2013. doi:10.1128/microbiolspec.MGM2-0024-2013.
- Gong, C., Bongiorno, P., Martins, A., Stephanou, N. C., Zhu, H., Shuman, S., *et al.* (2005). Mechanism of nonhomologous end-joining in mycobacteria: a low-fidelity repair system driven by Ku, ligase D and ligase C. *Nat. Struct. Mol. Biol.* 12, 304–312. doi:10.1038/nsmb915.
- Gong, C., Martins, A., Bongiorno, P., Glickman, M., and Shuman, S. (2004). Biochemical and genetic analysis of the four DNA ligases of mycobacteria. *J. Biol. Chem.* 279, 20594–20606. doi:10.1074/jbc.M401841200.

- Gordon, J. L., Byrne, K. P., and Wolfe, K. H. (2011). Mechanisms of chromosome number evolution in yeast. *PLoS Genet.* 7, e1002190.
- Gupta, R., Barkan, D., Redelman-Sidi, G., Shuman, S., and Glickman, M. S. (2011). Mycobacteria exploit three genetically distinct DNA double-strand break repair pathways. *Mol. Microbiol.* 79, 316–330. doi:10.1111/j.1365-2958.2010.07463.x.
- Gust, B., Challis, G. L., Fowler, K., Kieser, T., and Chater, K. F. (2003). PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proc. Natl. Acad. Sci. U. S. A.* 100, 1541–1546.
- Hoff, G., Bertrand, C., Piotrowski, E., Thibessard, A., and Leblond, P. (2016). Implication of RuvABC and RecG in homologous recombination in *Streptomyces ambofaciens*. *Res. Microbiol.* doi:10.1016/j.resmic.2016.07.003.
- Huang, T.-W., and Chen, C. W. (2006). A *recA* null mutation may be generated in *Streptomyces coelicolor*. *J. Bacteriol.* 188, 6771–6779. doi:10.1128/JB.00951-06.
- Inoue, S., Higashiyama, K., Uchida, T., Hiratsu, K., and Kinashi, H. (2003) Chromosomal circularization in *Streptomyces griseus* by nonhomologous recombination of deletion ends. *Biosci. Biotechnol. Biochem.* 67(5), 1101-1108.
- Ikeda, H., Ishikawa, J., Hanamoto, A., Shinose, M., Kikuchi, H., Shiba, T., *et al.* (2003). Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. *Nat. Biotechnol.* 21, 526–531. doi:10.1038/nbt820.
- Jacobs, M. A., Alwood, A., Thaipisuttikul, I., Spencer, D., Haugen, E., Ernst, S., *et al.* (2003). Comprehensive transposon mutant library of *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. U. S. A.* 100, 14339–14344.
- Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F., and Hopwood, D. A. (2000). *Practical Streptomyces genetics*. Norwich: John Innes Foundation.
- Kim, J.-N., Kim, Y., Jeong, Y., Roe, J.-H., Kim, B.-G., and Cho, B.-K. (2015). Comparative genomics reveals the core and accessory genomes of *Streptomyces* species. *J. Microbiol. Biotechnol.* 25, 1599–1605. doi:10.4014/jmb.1504.04008.
- Kobayashi, H., Simmons, L. A., Yuan, D. S., Broughton, W. J., and Walker, G. C. (2008). Multiple Ku orthologues mediate DNA non-homologous end-joining in the free-living form and during chronic infection of *Sinorhizobium meliloti*. *Mol. Microbiol.* 67, 350–363. doi:10.1111/j.1365-2958.2007.06036.x.
- Kushwaha, A. K., and Grove, A. (2013). C-terminal low-complexity sequence repeats of *Mycobacterium smegmatis* Ku modulate DNA binding. *Biosci. Rep.* 33. doi:10.1042/BSR20120105.

- Leblond, P., Demuyter, P., Simonet, J. M., and Decaris, B. (1990). Genetic instability and hypervariability in *Streptomyces ambofaciens*: towards an understanding of a mechanism of genome plasticity. *Mol. Microbiol.* 4, 707–714.
- Lin, Y.-S., Kieser, H. M., Hopwood, D. A., and Chen, C. W. (1993). The chromosomal DNA of *Streptomyces lividans* 66 is linear. *Mol. Microbiol.* 10, 923–933.
- Martínez-Burgo, Y., Álvarez-Álvarez, R., Rodríguez-García, A., and Liras, P. (2015). The pathway-specific regulator ClaR of *Streptomyces clavuligerus* has a global effect on the expression of genes for secondary metabolism and differentiation. *Appl. Environ. Microbiol.* 81, 6637–6648. doi:10.1128/AEM.00916-15.
- McGovern, S., Baconnais, S., Roblin, P., Nicolas, P., Drevet, P., Simonson, H., *et al.* (2016). C-terminal region of bacterial Ku controls DNA bridging, DNA threading and recruitment of DNA ligase D for double strand breaks repair. *Nucleic Acids Res.* 44, 4785–4806. doi:10.1093/nar/gkw149.
- Moeller, R., Stackebrandt, E., Reitz, G., Berger, T., Rettberg, P., Doherty, A. J., *et al.* (2007). Role of DNA repair by nonhomologous-end joining in *Bacillus subtilis* spore resistance to extreme dryness, mono- and polychromatic UV, and ionizing radiation. *J. Bacteriol.* 189, 3306–3311. doi:10.1128/JB.00018-07.
- Muth, G., Frese, D., Kleber, A., and Wohlleben, W. (1997). Mutational analysis of the *Streptomyces lividans recA* gene suggests that only mutants with residual activity remain viable. *Mol. Gen. Genet.* 255, 420–428.
- de Ory, A., Zafra, O., and de Vega, M. (2014). Efficient processing of abasic sites by bacterial nonhomologous end-joining Ku proteins. *Nucleic Acids Res.* 42, 13082–13095. doi:10.1093/nar/gku1029.
- de Ory, A., Nagler, K., Carrasco, B., Raguse, M., Zafra, O., Moeller, R. *et al.* (2016). Identification of a conserved 5'-dRP lyase activity in bacterial DNA repair ligase D and its potential role in base excision repair. *Nucleic Acids Res.* 44, 1833–1844. doi:10.1093/nar/gkw054.
- Paris, Ü., Mikkel, K., Tavita, K., Saumaa, S., Teras, R., and Kivisaar, M. (2015). NHEJ enzymes LigD and Ku participate in stationary-phase mutagenesis in *Pseudomonas putida*. *DNA Repair* 31, 11–18. doi:10.1016/j.dnarep.2015.04.005.
- Pinnert-Sindico, S. (1954). Une nouvelle espèce de *Streptomyces* productrice d'antibiotiques: *Streptomyces ambofaciens* n. sp. caractères cultureux. *Ann Inst Pasteur.* 87, 702-707.
- Pitcher, R. S., Green, A. J., Brzostek, A., Korycka-Machala, M., Dziadek, J., and Doherty, A. J. (2007). NHEJ protects mycobacteria in stationary phase against the harmful effects of desiccation. *DNA Repair* 6, 1271–1276. doi:10.1016/j.dnarep.2007.02.009.

- Pitcher, R. S., Tonkin, L. M., Green, A. J., and Doherty, A. J. (2005). Domain structure of a NHEJ DNA repair ligase from *Mycobacterium tuberculosis*. *J. Mol. Biol.* 351, 531–544. doi:10.1016/j.jmb.2005.06.038.
- Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 29(9), 2002–2007
- Raynal, A., Karray, F., Tuphile, K., Darbon-Rongere, E., and Pernodet, J.-L. (2006). Excisable cassettes: new tools for functional analysis of *Streptomyces* genomes. *Appl. Environ. Microbiol.* 72, 4839–4844. doi:10.1128/AEM.00167-06.
- Rocha, E. P. C., Cornet, E., and Michel, B. (2005). Comparative and evolutionary analysis of the bacterial homologous recombination systems. *PLoS Genet.* 1, e15. doi:10.1371/journal.pgen.0010015.
- Seipke, R. F., Kaltenpoth, M., and Hutchings, M. I. (2012). *Streptomyces* as symbionts: an emerging and widespread theme? *FEMS Microbiol. Rev.* 36, 862–876. doi:10.1111/j.1574-6976.2011.00313.x.
- Shee, C., Gibson, J. L., and Rosenberg, S. M. (2012). Two mechanisms produce mutation hotspots at DNA breaks in *Escherichia coli*. *Cell Rep.* 2, 714–721. doi:10.1016/j.celrep.2012.08.033.
- Stephanou, N. C., Gao, F., Bongiorno, P., Ehrt, S., Schnappinger, D., Shuman, S., *et al.* (2007). Mycobacterial nonhomologous end joining mediates mutagenic repair of chromosomal double-strand DNA breaks. *J. Bacteriol.* 189, 5237–5246. doi:10.1128/JB.00332-07.
- Sun, J., Kelemen, G. H., Fernández-Abalos, J. M., and Bibb, M. J. (1999). Green fluorescent protein as a reporter for spatial and temporal gene expression in *Streptomyces coelicolor* A3 (2). *Microbiol.* 145, 2221–2227.
- Tamura, K., Stecher, G., Peterson, D., Filipowski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi:10.1093/molbev/mst197.
- Thibessard, A., Haas, D., Gerbaud, C., Aigle, B., Lautru, S., Pernodet, J.-L., *et al.* (2015). Complete genome sequence of *Streptomyces ambofaciens* ATCC 23877, the spiramycin producer. *J. Biotechnol.* 214, 117–118. doi:10.1016/j.jbiotec.2015.09.020.
- Thibessard, A., and Leblond, P. (2014) “Subtelomere plasticity in the bacterium *Streptomyces*,” in *Subtelomeres*, ed. Springer, 243-258.
- Uchida, T., Miyawaki, M., and Kinashi, H. (2003). Chromosomal arm replacement in *Streptomyces griseus*. *J. Bacteriol.* 185, 1120–1124. doi:10.1128/JB.185.3.1120-1124.2003.

- Vierling, S., Weber, T., Wohlleben, W., and Muth, G. (2001). Evidence that an additional mutation is required to tolerate insertional inactivation of the *Streptomyces lividans recA* gene. *J. Bacteriol.* 183, 4374–4381. doi:10.1128/JB.183.14.4374-4381.2001.
- Volff, J.-N., and Altenbuchner, J. (1998). Genetic instability of the *Streptomyces* chromosome. *Mol. Microbiol.* 27, 239–246.
- Wang, S. T., Setlow, B., Conlon, E. M., Lyon, J. L., Imamura, D., Sato, T., *et al.* (2006). The forespore line of gene expression in *Bacillus subtilis*. *J. Mol. Biol.* 358, 16–37. doi:10.1016/j.jmb.2006.01.059.
- Weller, G. R., Kysela B., Roy R., Tonkin L. M., Scanlan E., Della M., *et al.* (2002). Identification of a DNA nonhomologous end-joining complex in bacteria. *Science* 297, 1686–1689. doi:10.1126/science.1074584.
- Wenner, T., Roth, V., Fischer, G., Fourrier, C., Aigle, B., Decaris, B., *et al.* (2003). End-to-end fusion of linear deleted chromosomes initiates a cycle of genome instability in *Streptomyces ambofaciens*. *Mol. Microbiol.* 50, 411–425. doi:10.1046/j.1365-2958.2003.03698.x.
- Yanai, K., Murakami, T., and Bibb, M. (2006). Amplification of the entire kanamycin biosynthetic gene cluster during empirical strain improvement of *Streptomyces kanamyceticus*. *Proc. Natl. Acad. Sci. U. S. A.* 103, 9661–9666.
- Zhang, L., Nguyen, H. C., Chipot, L., Piotrowski, E., Bertrand, C., Thibessard, A., *et al.* (2014). The *adnAB* locus, encoding a putative helicase-nuclease activity, is essential in *Streptomyces*. *J. Bacteriol.* 196, 2701–2708. doi:10.1128/JB.01513-14.
- Zhang, X., Chen, W., Zhang, Y., Jiang, L., Chen, Z., Wen, Y., *et al.* (2012). Deletion of *ku* homologs increases gene targeting frequency in *Streptomyces avermitilis*. *J. Ind. Microbiol. Biotechnol.* 39, 917–925. doi:10.1007/s10295-012-1097-x.
- Zhu, H., Bhattarai, H., Yan, H.-G., Shuman, S., and Glickman, M. S. (2012). Characterization of *Mycobacterium smegmatis* PolD2 and PolD1 as RNA/DNA polymerases homologous to the POL domain of bacterial DNA ligase D. *Biochem.* 51, 10147–10158. doi:10.1021/bi301202e.
- Zhu, H., and Shuman, S. (2007). Characterization of *Agrobacterium tumefaciens* DNA ligases C and D. *Nucleic Acids Res.* 35, 3631–3645. doi:10.1093/nar/gkm145.
- Zhu, H., and Shuman, S. (2010). Gap filling activities of *Pseudomonas* DNA ligase D (LigD) polymerase and functional interactions of LigD with the DNA end-binding Ku protein. *J. Biol. Chem.* 285, 4815–4825. doi:10.1074/jbc.M109.073874.

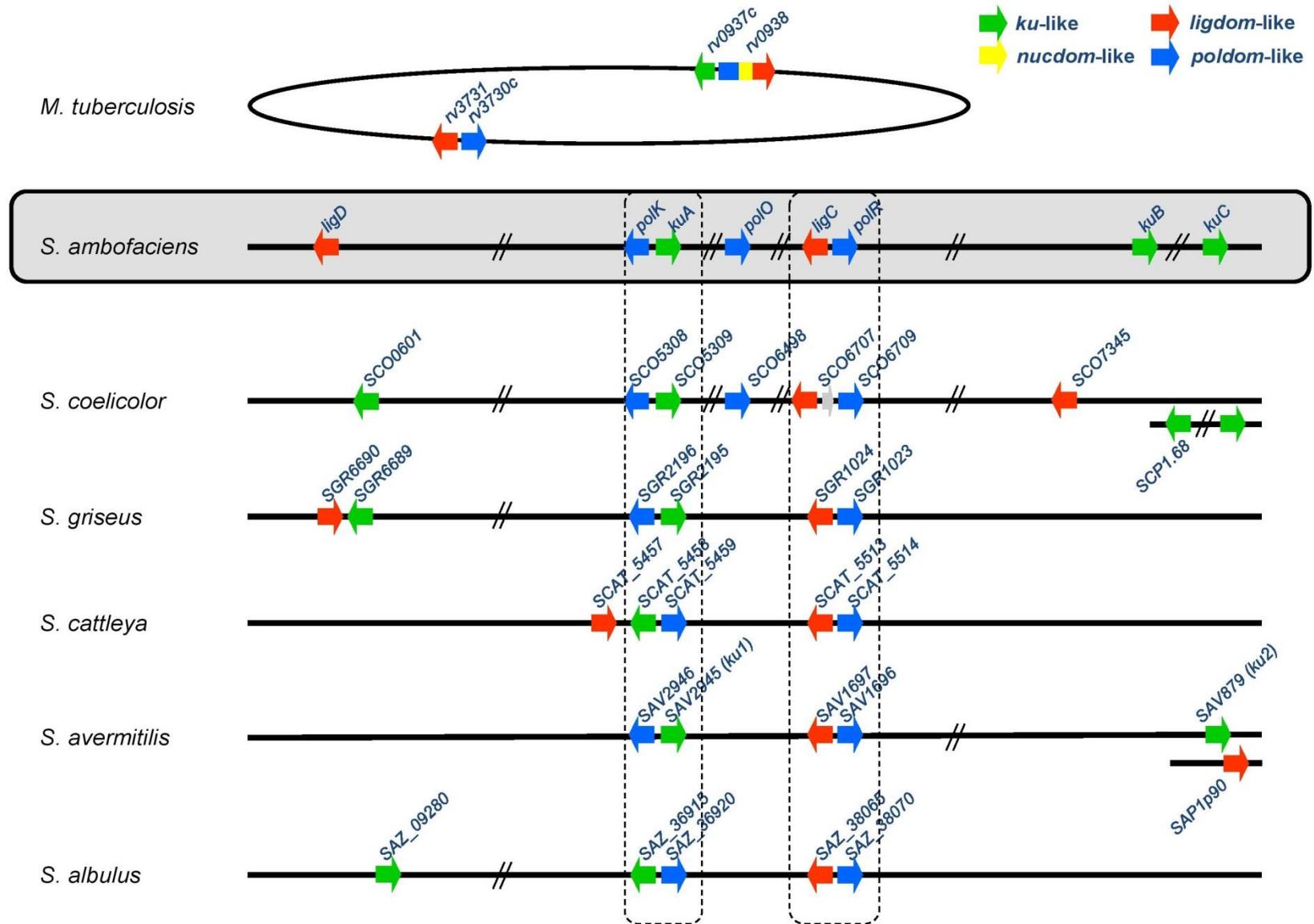


Figure 1: Organization of the NHEJ-like genes in several *Streptomyces* genomes. The two conserved loci *kuA-polK* and *ligC-polR* are framed in dotted lines.

Organism name	Strain	Size (Mb)	kuA	kuB	kuC	ligC	ligD	polK	polO	polR
<i>S. xiamenensis</i>	318	5.96								
<i>S. albus</i>	J1074	6.84								
<i>S. sp. Tu6071</i>	Tu6071	7.36			*					
<i>S. sp. SirexAA-E</i>	SirexAA-E	7.41					**			
<i>S. glaucescens</i>	GLA.O	7.62								
<i>S. sp. PAMC26508</i>	PAMC26508	7.63					**			
<i>S. pratensis</i>	ATCC 33331	7.66					**			
<i>S. cyaneogriseus</i> subsp. <i>noncyanogenus</i>	NMWT 1	7.76								
<i>S. globisporus</i>	C-1027	7.78					**			
<i>S. fulvissimus</i>	DSM 40593	7.91					***			
<i>S. sp. CNQ-509</i>	CNQ-509	8.04								
<i>S. sp. 4F</i>	4F	8.05								
<i>S. cattleya</i>	NRRL 8057 ^a	8.09								
<i>S. leeuwenhoekii</i>	C34 ^b	8.12			*					
<i>S. sp. CFMR 7</i>	CFMR-7	8.31			*		*			
<i>S. lividans</i>	TK24	8.35								
<i>S. collinus</i>	Tu 365	8.38								
<i>S. ambofaciens</i>	ATCC 23877	8.39								
<i>S. pristinaespiralis</i>	HCCB 10218	8.53								
<i>S. griseus</i> subsp. <i>griseus</i>	NBRC 13350	8.55					**			
<i>S. coelicolor</i>	A3(2)	8.67			2 *					
<i>S. sp. Mg1</i>	Mg1	8.72		*						
<i>S. venezuelae</i>	ATCC 15439	9.03								
<i>S. avermitilis</i>	MA-4680 ^c	9.12								
<i>S. vietnamensis</i>	GIM4.0001	9.15								
<i>S. lydicus</i>	A02	9.30	2							
<i>S. reticuli</i>	TUE45	9.34			*					
<i>S. albulus</i>	NK660	9.37	2							
<i>S. davawensis</i>	JCM4913	9.47								
<i>S. albulus</i>	ZPM	9.78	2							
<i>S. hygrosopicus</i> subsp. <i>jinggangensis</i>	TL01	10.08			*					
<i>S. scabiei</i>	87.22	10.15								
<i>S. sp. CdTB01</i>	CdTB01	10.19								
<i>S. sp. 769</i>	769	10.34	2							
<i>S. hygrosopicus</i> subsp. <i>jinggangensis</i>	5008	10.38			*					
<i>S. hygrosopicus</i> subsp. <i>limoneus</i>	KCTC 1717	10.54								
<i>S. violaceusniger</i>	Tu 4113	11.14	3							
<i>S. bingchengensis</i>	BCW-1	11.94								

^a NRRL 8057 = DSM 46488; ^b C34 = DSM 42122 = NRRL B-24963; ^c MA-4680 (= NBRC 14893)

* plasmidic

** closed to *kuC*

*** partial (35% of cover)

Figure 2: Distribution of NHEJ-like genes in 38 complete genomes of *Streptomyces*. The presence of a homologue is mentioned by a black box and the absence of a homolog by an empty box. The number inside the boxes indicates when several homologues were identified; * plasmidic; ** close to *kuC*; *** partial (cover of 35%).

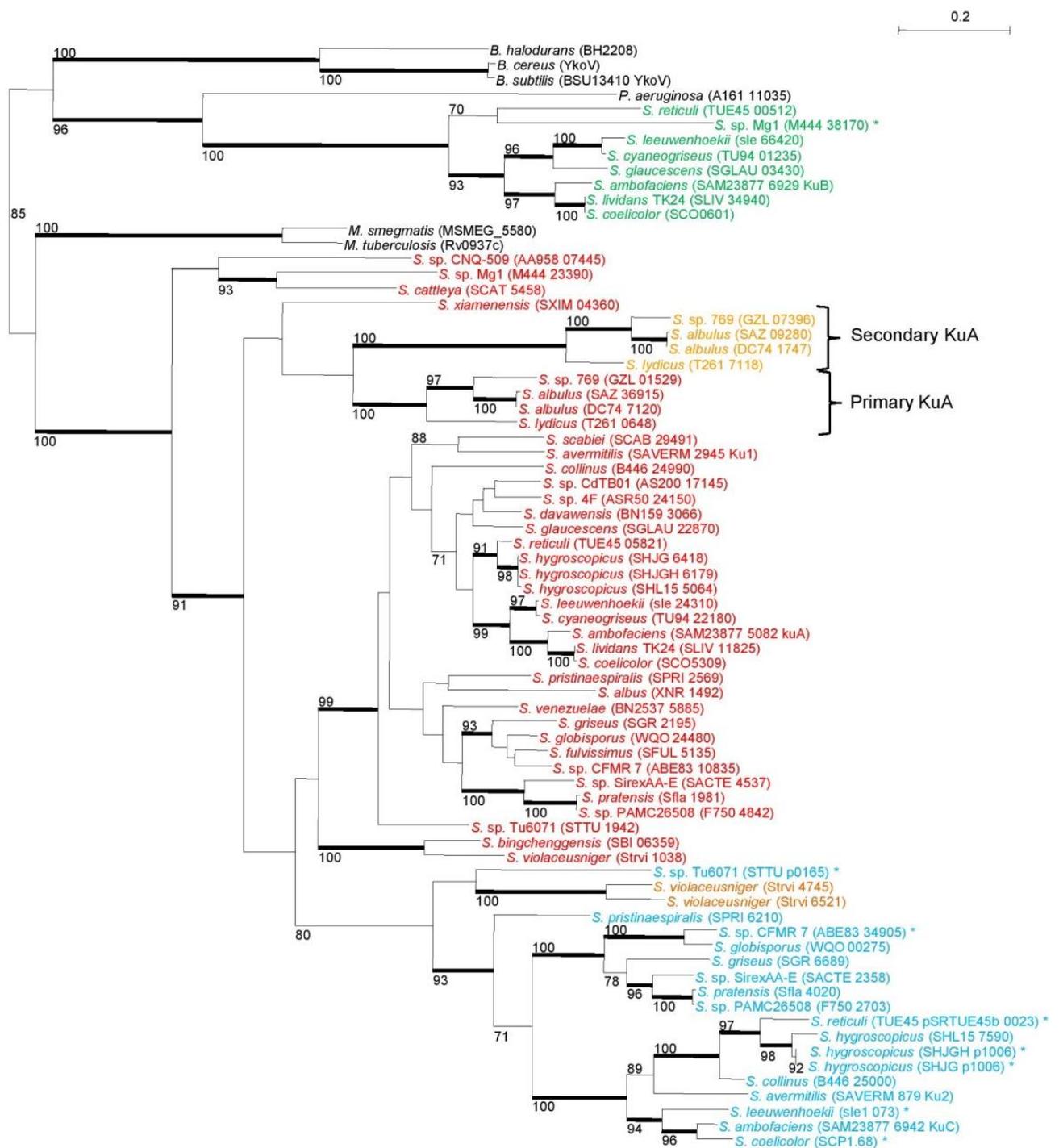


Figure 3: Phylogeny of Ku-like proteins identified in 38 *Streptomyces* strains. *Streptomyces* Ku-like proteins were identified in 38 *Streptomyces* strains and compared with the closest homologues of *M. tuberculosis*, *M. smegmatis*, *B. subtilis*, *B. cereus*, *B. halodurans* and *P. aeruginosa*. The tree was built using a maximum-likelihood method with a 295 aa alignment. Only bootstrap values higher than 70% are reported. Heavy lines indicate branches supported by bootstrap values > 90% (100 replicates). The scale represents mutations per amino-acid. Primary KuA are in red, secondary KuA are in orange, KuB are in green and KuC in blue. The asterisks symbolize the proteins encoded by plasmid borne-genes. The locus-tag of each sequence is mentioned in brackets. The suspected event of gene duplication is depicted by the two braces. B : *Bacillus*, M : *Mycobacterium*, P : *Pseudomonas*, S : *Streptomyces*. ^a NRRL 8057 = DSM 46488; ^b C34 = DSM 42122 = NRRL B-24963; ^c MA-4680 (= NBRC 14893).

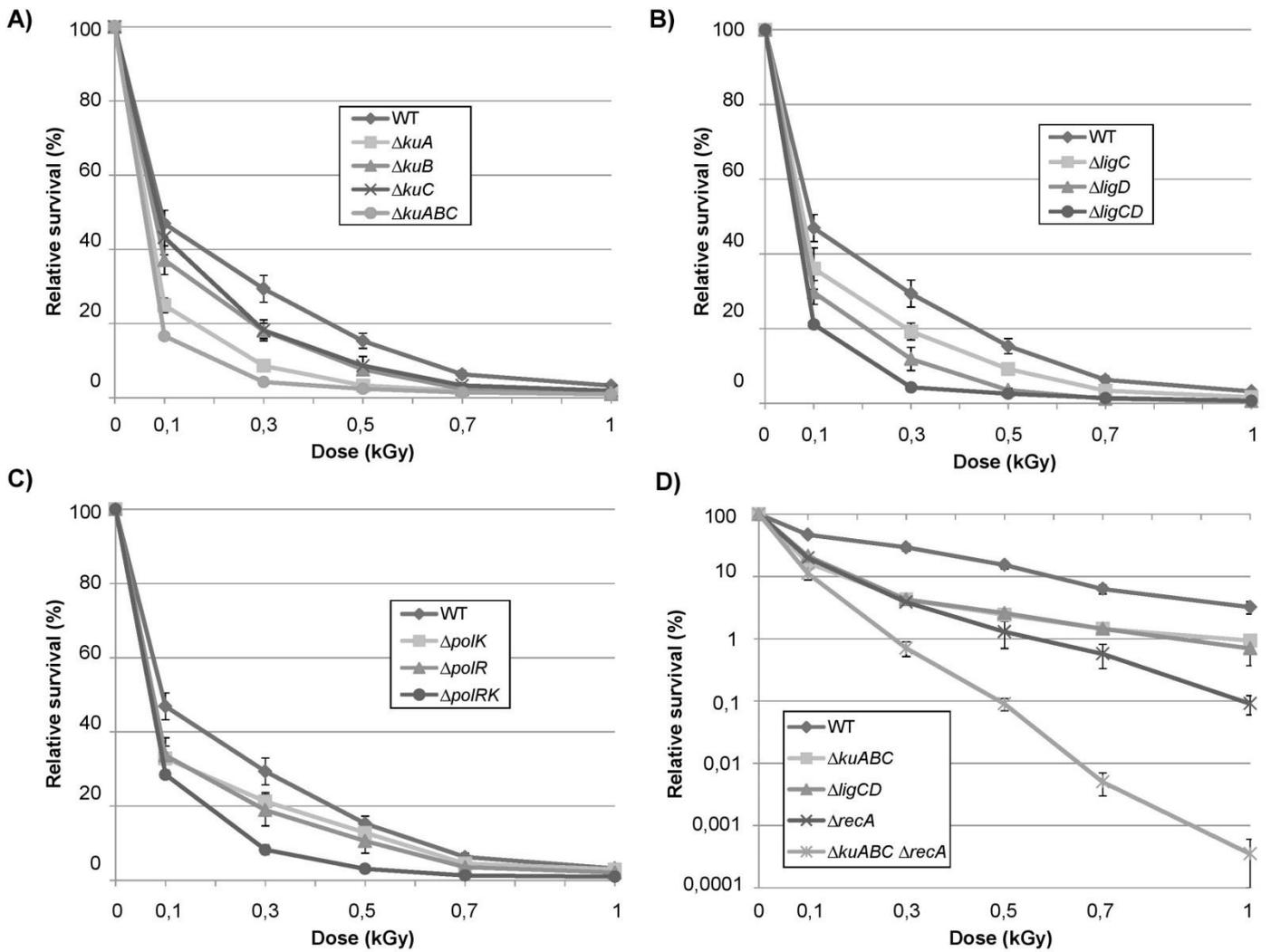


Figure 4: Electronic beam irradiation of *S. ambofaciens* strains deficient in NHEJ-like genes. Spores of mutants of NHEJ-like genes, namely single and multiple *ku* (A), ligase (B) and polymerase (C) mutants, were exposed to 0.1, 0.3, 0.5, 0.7 and 1 kGy of electronic beam irradiation. (D) In the same way, a Δku -like $\Delta recA$ mutant also affected in HR was exposed to the same irradiation doses. In this latter case, a logarithmic scale was used for survival rate representation. Survival rates were calculated as the number of clones for a given dose relative to the number of clones unexposed to radiation. For each graph, wild type strain is represented as a control. Error bars indicate standard error.

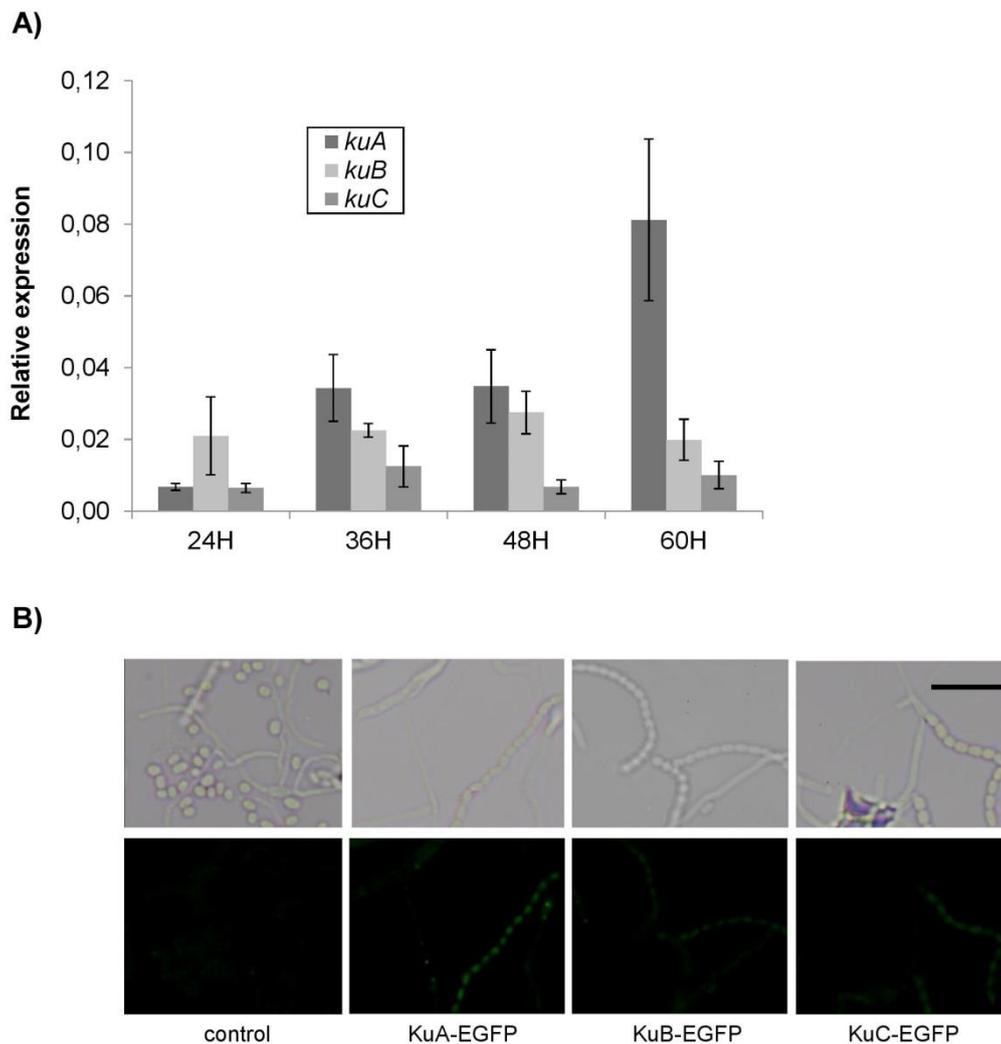


Figure 5: Growth-dependent and cell-type specific expression of NHEJ-like genes. (A) Expression level of *S. ambofaciens* ATCC 23877 *ku* genes relative to *hrdB* after 24, 36, 48 and 60 hours growth on SFM plates at 30°C. Error bars represent standard error. **(B)** Strains expressing EGFP fusion with KuA, KuB or KuC were inoculated in the acute-angle junction of standard-sized microscope coverslips inserted at 45° in SFM agar, and grown for 48 hours before observations by fluorescent microscopy. The control strain does not contain any GFP gene. Representative examples of spores or aerial mycelium are shown as a contrast phase image (upper panels) and as EGFP fluorescent channel (lower panels). The weak fluorescent foci observed mainly in mycelia were readily detected in the control strain and are not likely to correspond to GFP signal. The scale is represented by a 5 µm size bar.

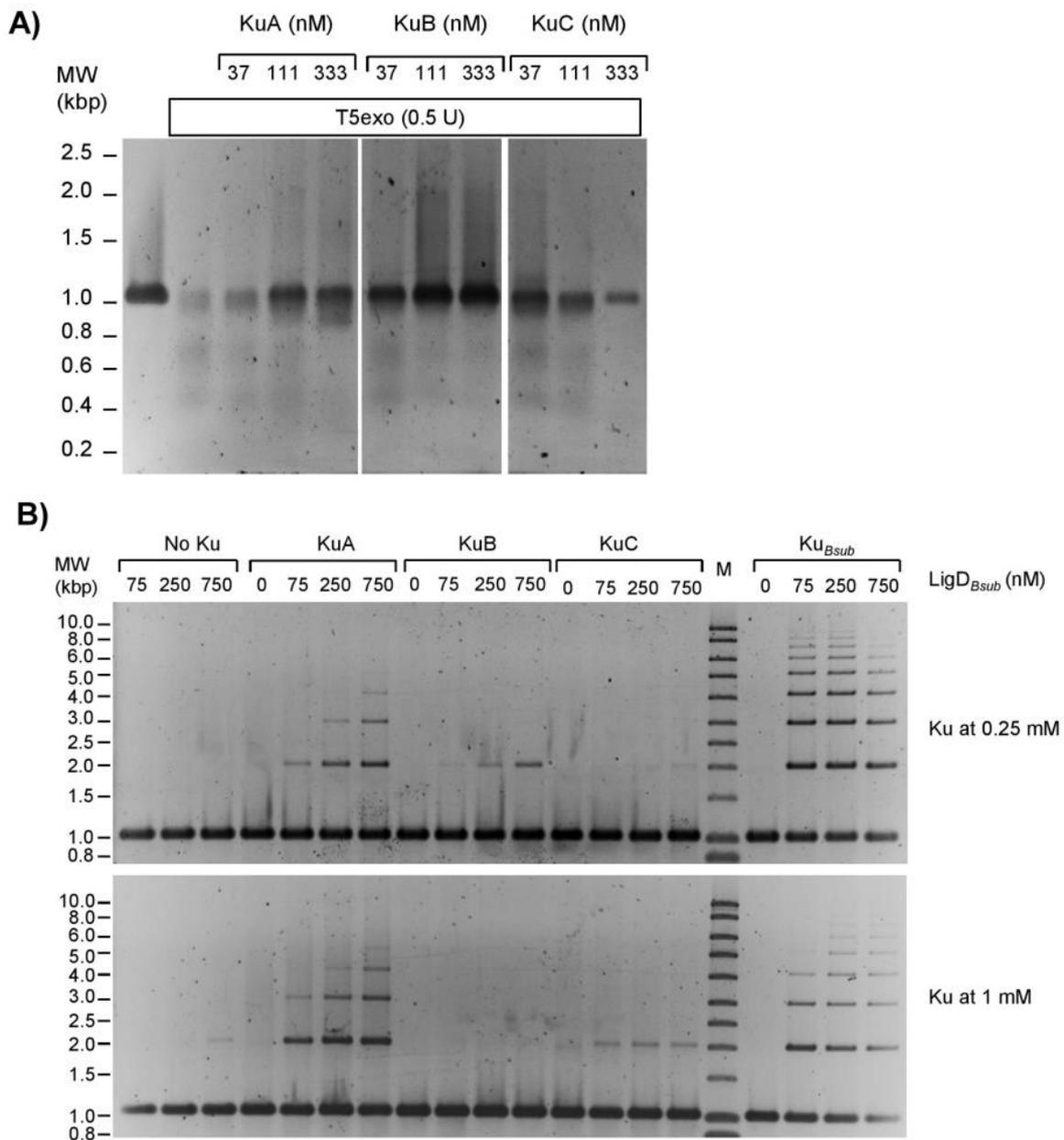


Figure 6: KuA, KuB and KuC protect linear dsDNA from the T5 exonuclease and stimulate the ligase activity of LigD_{Bsub}. (A) A 1001 bp linear dsDNA with 5'-phosphorylated ends (3.8 nM) was incubated 30 min at 37°C with increasing amounts of KuA, KuB or KuC as indicated. T5 exonuclease (0.5 U) was then added in each reaction mixtures for 30 minutes at 37°C and digested products were analyzed as described in the Material and Methods. Control experiments showing the 1001 ds DNA substrate and the nearly complete degradation of the DNA molecule by the T5 exonuclease without other proteins added are shown on the left. (B) A 1001 bp linear dsDNA with 5'-phosphorylated ends (3.8 nM) was incubated with 0.25 or 1 μM of KuA, KuB or KuC as indicated. Increasing amounts of LigD_{Bsub} were then added (allowing final concentrations of LigD_{Bsub} indicated above the gels) and ligation products were analyzed as described in the Material and Methods. Control experiments showing the weak activity of LigD_{Bsub} alone at different concentration are shown on the left (No Ku). For comparison, Ku_{Bsub} was assessed for its LigD_{Bsub} stimulation activity in the same conditions (right part).

TABLE 1 | *Streptomyces ambofaciens* strains used in this study.

Strains	Characteristics	Sources
ATCC23877	Used as a WT strain	Pinnert-Sindico, 1954
ΔkuA	$\Delta SAM23877_5082$	This study
ΔkuB	$\Delta SAM23877_6929$	This study
ΔkuC	$\Delta SAM23877_6942$	This study
$\Delta kuABC$	$\Delta SAM23877_5082$ $\Delta SAM23877_6929$ $\Delta SAM23877_6942$	This study
$\Delta ligC$	$\Delta SAM23877_6361$	This study
$\Delta ligD$	$\Delta SAM23877_0862$	This study
$\Delta ligCD$	$\Delta SAM23877_6361$ $\Delta SAM23877_0862$	This study
$\Delta polR$	$\Delta SAM23877_6362$	This study
$\Delta polK$	$\Delta SAM23877_5081$	This study
$\Delta polRK$	$\Delta SAM23877_6362$ $\Delta SAM23877_5081$	This study
$\Delta recA$	$\Delta SAM23877_5473$	This study
$\Delta kuABC \Delta recA$	$\Delta SAM23877_5082$ $\Delta SAM23877_6929$ $\Delta SAM23877_6942$ $\Delta SAM23877_5473$	This study

SUPPLEMENTARY MATERIAL

Gel filtration assay

Two nmol of KuA, KuB or KuC proteins were incubated in 500 μ L of buffer G (50 mM Tris-HCl, pH 8, 250 mM NaCl) for 10 minutes at 4°C. After centrifugation at 13000 rpm, the soluble fraction was injected on a Superdex 200 10/300 GL (GE Healthcare) equilibrated with buffer G. Twenty two fractions of 500 μ L were collected, ranging from approximately 700 kDa (the void volume) to 10 kDa. To compare the elution volumes of the proteins, 5 μ L of SDS-loading buffer were added to 20 μ L of each of these fractions and loaded onto 10 % acrylamide SDS-PAGE gels. After migration, gels were stained with Coomassie blue.

Table S1: Identification of NHEJ-like genes in 38 complete genomes of *Streptomyces* by their locus_tag

Organism ^a	Strain	Ku-like			ATP-dependant DNA ligase		Poldom-like		
		KuA	KuB	KuC	LigC	LigD	PolK	PolO	PolR
<i>S. xiamenensis</i>	318	SXIM_04360	-	-	SXIM_50860	SXIM_51150	SXIM_04370	-	SXIM_50850
<i>S. albus</i>	J1074	XNR_1492	-	-	XNR_0334	-	XNR_1493	XNR_4488	XNR_0333
<i>S. sp. Tu6071</i>	Tu6071	STTU_1942	-	STTU_p0165 (plasmidic)	STTU_0771	-	STTU_1943	-	STTU_0755
<i>S. sp. SirexAA-E</i>	SirexAA-E	SACTE_4537	-	SACTE_2358	SACTE_5876	SACTE_2357 **	SACTE_4536	-	SACTE_5877
<i>S. glaucescens</i>	GLA.O	SGLAU_22870	SGLAU_03430	-	SGLAU_28040	SGLAU_02470	SGLAU_22865	SGLAU_27360	SGLAU_28045
<i>S. sp. PAMC26508</i>	PAMC26508	F750_4842	-	F750_2703	F750_6167	F750_2702 **	F750_4841	-	F750_6168
<i>S. pratensis</i>	ATCC 33331	Sfla_1981	-	Sfla_4020	Sfla_0697	Sfla_4021 **	Sfla_1982	-	Sfla_0696
<i>S. cyaneogriseus</i>	NMWT 1	TU94_22180	TU94_01235	-	TU94_28785	TU94_01705	TU94_22175	-	TU94_28790
<i>S. globisporus</i>	C-1027	WQO_24480	-	WQO_00275	WQO_30485	WQO_00280 **	WQO_24475	-	WQO_30490
<i>S. fulvissimus</i>	DSM 40593	SFUL_5135	-	-	SFUL_6473	SFUL_6958 partial	SFUL_5134	-	SFUL_6474
<i>S. sp. CNQ-509</i>	CNQ-509	AA958_07445	-	-	AA958_31995	AA958_07435	AA958_07440	-	AA958_31990
<i>S. sp. 4F</i>	4F	ASR50_24150	-	-	ASR50_30225	ASR50_04130	ASR50_24145	ASR50_29170	ASR50_30230
<i>S. cattleya</i>	NRRL 8057	SCAT_5458	-	-	SCAT_5513	SCAT_5457	SCAT_5459	-	SCAT_5514
<i>S. leeuwenhoekii</i>	C34	sle_24310	sle_66420	sle1_073 (plasmidic)	sle_11180	sle_65450	sle_24320	sle_12820	sle_11170
<i>S. sp. CFMR 7</i>	CFMR-7	ABE83_10835	-	ABE83_34905 (plasmidic)	ABE83_03950	ABE83_34900 (plasmidic) **	ABE83_10840	-	ABE83_03945
<i>S. lividans</i>	TK24	SLIV_11825	SLIV_34940	-	SLIV_04970	SLIV_02570	SLIV_11830	SLIV_05935	SLIV_04965
<i>S. collinus</i>	Tu 365	B446_24990	-	B446_25000	B446_30620	B446_04080	B446_24985	-	B446_30625
<i>S. ambofaciens</i>	ATCC 23877	SAM23877_5082	SAM23877_6929	SAM23877_6942	SAM23877_6361	SAM23877_0862	SAM23877_5081	SAM23877_6202	SAM23877_6362
<i>S. pristinaespiralis</i>	HCCB 10218	SPRI_2569	-	SPRI_6210	SPRI_1068	-	SPRI_2570	SPRI_6912	SPRI_1067
<i>S. griseus</i>	NBRC 13350	SGR_2195	-	SGR_6689	SGR_1024	SGR_6690 **	SGR_2196	-	SGR_1023
<i>S. coelicolor</i>	A3(2)	SCO5309	SCO0601	SCP1.68 (plasmidic)	SCO6707	SCO7345	SCO5308	SCO6498	SCO6709
<i>S. sp. Mg1</i>	Mg1	M444_23390	M444_38170 (plasmidic)	-	M444_27900	-	M444_23395	-	M444_27905
<i>S. venezuelae</i>	ATCC 15439	BN2537_5885	-	-	BN2537_3729	-	BN2537_5887	-	BN2537_3727
<i>S. avermitilis</i>	MA-4680	SAVERM_2945	-	SAVERM_879	SAVERM_1697	-	SAVERM_2946	-	SAVERM_1696
<i>S. vietnamensis</i>	GIM4.0001	-	-	-	SVTN_31170	-	-	SVTN_30635	SVTN_31290
<i>S. lydicus</i>	A02	T261_0648 ; T261_7118	-	-	T261_0463	T261_1796	T261_0647	T261_1305	T261_0462
<i>S. reticuli</i>	-	TUE45_05821	TUE45_00512	pSRTUE45b_0023 (plasmidic)	TUE45_07256	TUE45_00866	TUE45_05820	TUE45_07099	TUE45_07257
<i>S. albulus</i>	NK660	DC74_7120 ; DC74_1747	-	-	DC74_7353	DC74_6447	DC74_7121	-	DC74_7354
<i>S. davawensis</i>	JCM4913	BN159_3066	-	-	BN159_1716	BN159_8336	BN159_3067	-	BN159_1715
<i>S. albulus</i>	ZPM	SAZ_36915 ; SAZ_09280	-	-	SAZ_38065	SAZ_33260	SAZ_36920	-	SAZ_38070
<i>S. hygroscopicus</i>	TL01	SHJGH_6179	-	SHJGH_p1006 (plasmidic)	SHJGH_7371	SHJGH_1851	SHJGH_6178	SHJGH_7216	SHJGH_7372
<i>S. scabiei</i>	87.22	SCAB_29491	-	-	SCAB_13591	-	SCAB_29521	SCAB_17401	SCAB_13581
<i>S. sp. CdTB01</i>	CdTB01	AS200_17145	-	-	AS200_09240	AS200_41235	AS200_17150	AS200_10695	AS200_09235
<i>S. sp. 769</i>	769	GZL_01529 ; GZL_07396	-	-	GZL_01249	GZL_02561	GZL_01528	-	GZL_01248
<i>S. hygroscopicus</i>	5008	SHJG_6418	-	SHJG_p1006 (plasmidic)	SHJG_7610	SHJG_2086	SHJG_6417	SHJG_7456	SHJG_7611
<i>S. hygroscopicus</i>	KCTC 1717	SHL15_5064	-	SHL15_7590	SHL15_6478	SHL15_0572	SHL15_5063	SHL15_6335	SHL15_6479
<i>S. violaceusniger</i>	Tu 4113	Strvi_1038 ; Strvi_6521 ; Strvi_4745	-	-	Strvi_3581	Strvi_5345	Strvi_1039	Strvi_0339	Strvi_3580
<i>S. bingchenggensis</i>	BCW-1	SBI_06359	-	-	SBI_08910	-	SBI_06360	-	SBI_08909

^a the species have been ordered by their genome size, from the smaller to the larger.

** *ligD* gene colocalizing with *kuC* gene

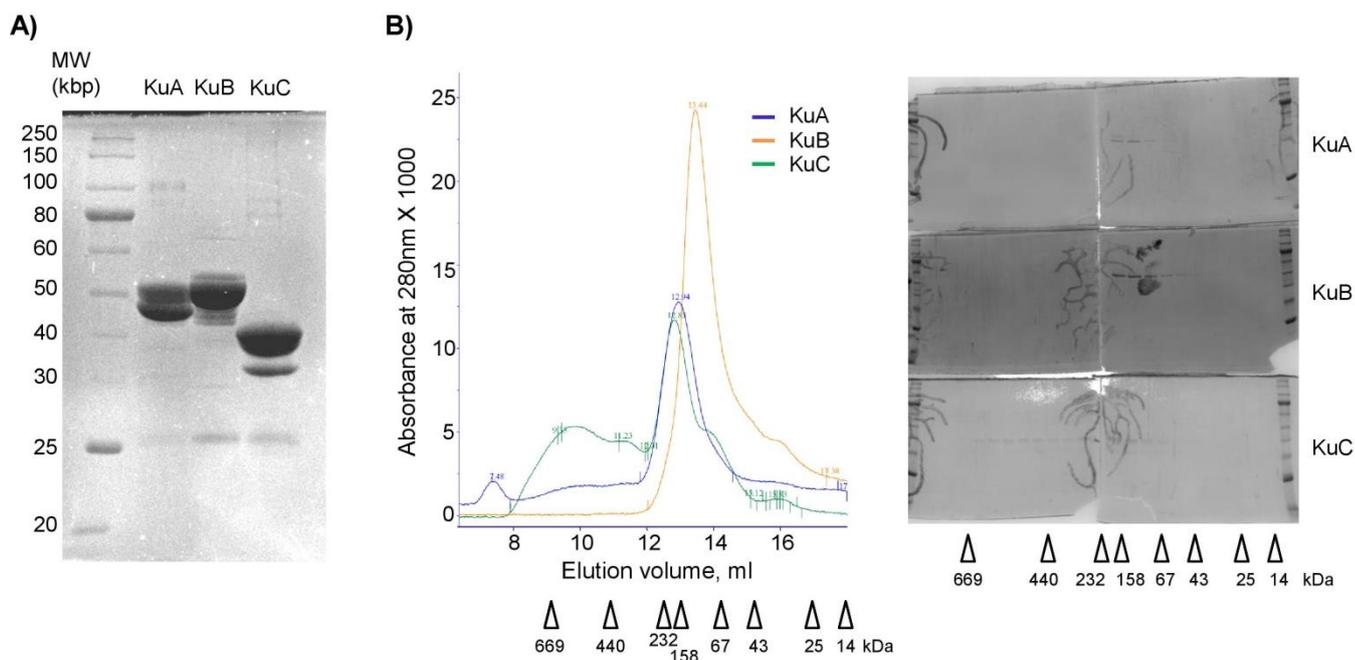


Figure S1: Analysis of the purified KuA, KuB and KuC proteins. A) SDS PAGE analysis of purified fractions containing 10 μ g of proteins obtained after elution from Ni-NTA agarose and dialysis. B) Gel filtration analysis of the purified Ku proteins. Chromatograms are presented on the left and protein contents analyzed by SDS PAGE on the right. Nearly all proteins in the KuA or KuB preparations were eluted in one elution peak corresponding to a molecular weight region around 160 kDa for KuA (39.1 kDa for the monomer) and 100 kDa for KuB (42.4 kDa for the monomer). Main proteins in the fraction containing KuC were co-eluted all along the gel filtration chromatography, from the void volume (molecular weight greater than 700 kDa) to elution volumes corresponding to 70 kDa (33.6 kDa for the monomer).

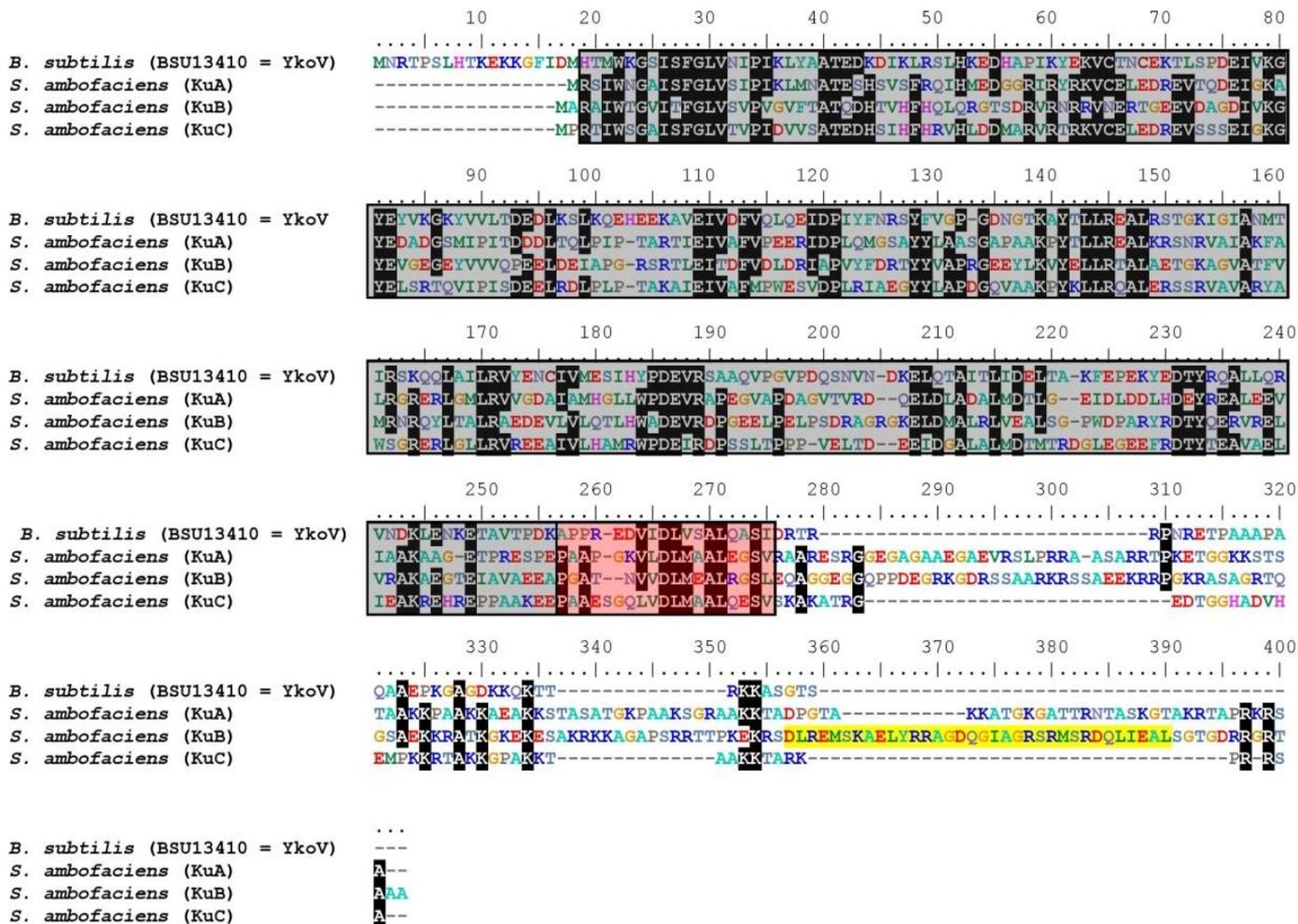


Figure S2: Alignment of amino-acid sequences of Ku-like proteins of *S. ambofaciens*. KuA, KuB and KuC proteins of *S. ambofaciens* composed of 365, 383 and 302 amino-acids respectively and the 311 amino-acid Ku_{Bsub} were aligned with ClustalW. The Ku core domain and the C-terminal minimal domain are boxed in grey and red, respectively. The SAP domain previously predicted in the *S. coelicolor* KuB homologue by (Aravind and Koonin 2001) is highlighted in yellow.

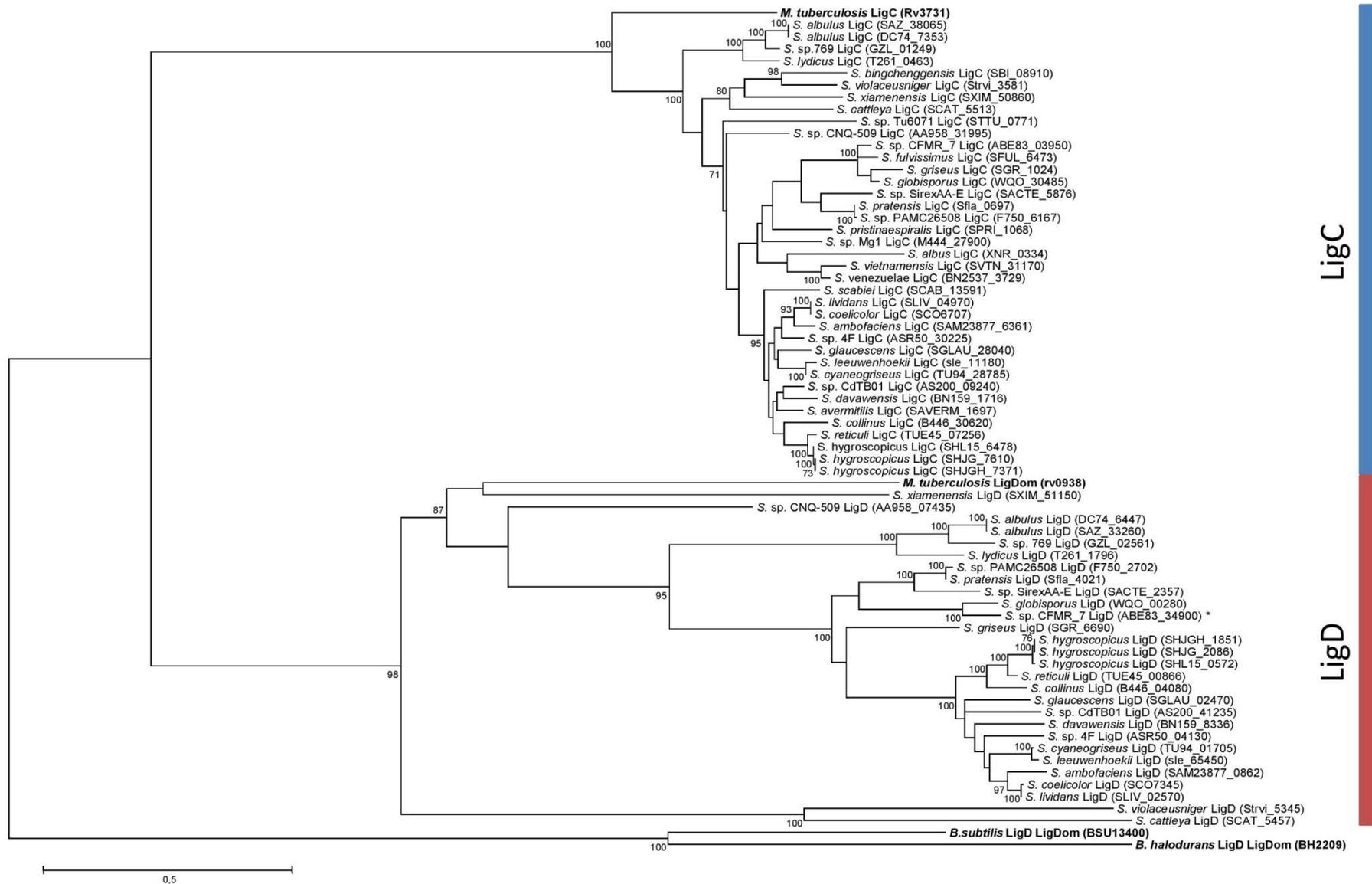


Figure S2. Phylogeny of *Streptomyces* LigC and LigD proteins. *Streptomyces* LigC and LigD proteins identified in 38 *Streptomyces* strains were compared with their homologues in *M. tuberculosis*. LigD homologues in *B. subtilis* and *B. halodurans* were used as a root. The tree was built using a maximum-likelihood method with a 309 aa position alignment. Numbers indicate by bootstrap values > 70% (100 replicates). The scale represents mutations per amino acid. The locus-tag of each sequence is mentioned in brackets. B : *Bacillus*, M : *Mycobacterium*, S : *Streptomyces*.

Résultats complémentaires :

Analyses microscopiques

Lors de la titration par spectrophotométrie de spores d'un mutant ΔkuA , nous avons suspecté une déficience de ce dernier dans la germination. Afin de vérifier l'intégrité des spores, nous avons entrepris des analyses microscopiques. Pour cela nous avons dans un premier temps réalisé sur du mycélium en phase de sporulation, des colorations au DAPI (4',6-diamidino-2-phénylindole), une molécule fluorescente qui se fixe aux bases A et T et qui permet de vérifier la présence de l'ADN au sein des cellules. L'objectif était de déterminer si le mutant ΔkuA produisait une proportion anormale de spores ne contenant pas de chromosome, appelées spores « fantômes ». Ces spores peuvent apparaître par exemple après un défaut de ségrégation des chromosomes lors de la septation de l'hyphe, et elles sont distinguables des spores normales car elles ne sont donc pas colorées au DAPI. Ces analyses ont révélé que les spores de la souche ΔkuA ne présentaient pas plus de spores fantômes que la souche sauvage. Par la suite, les spores ont également été observées au microscope électronique à balayage (MEB) ainsi qu'au microscope électronique à transmission (MET). Pour cela, du mycélium en phase de sporulation a été prélevé d'une souche WT et d'une souche ΔkuA puis analysé. Les résultats des observations de chaînes de spores au MEB et au MET sont présentés en **figure R1**.

Les images obtenues au MEB nous ont permis de voir que l'ensemble des hyphes observés dans les deux contextes génétiques étaient en phase de septation, et produisaient des chaînes de spores continues, avec des spores de taille homogène (**figure R1A**). De la même manière, les images obtenues au MET nous ont permis de voir que la souche sauvage et la souche ΔkuA produisaient des spores matures (**figure R1B**). Ainsi les observations microscopiques n'ont révélé aucune différence au niveau des spores entre la souche sauvage et la souche ΔkuA . L'analyse de structure des spores ne nous a pas permis de déterminer ce qui pourrait provoquer une déficience de germination. Par ailleurs, après avoir essayé différentes techniques peu concluantes (cellule de Mallassez, cytométrie de flux), nous n'avons pas pu démontrer que le mutant ΔkuA avait effectivement un problème de germination.

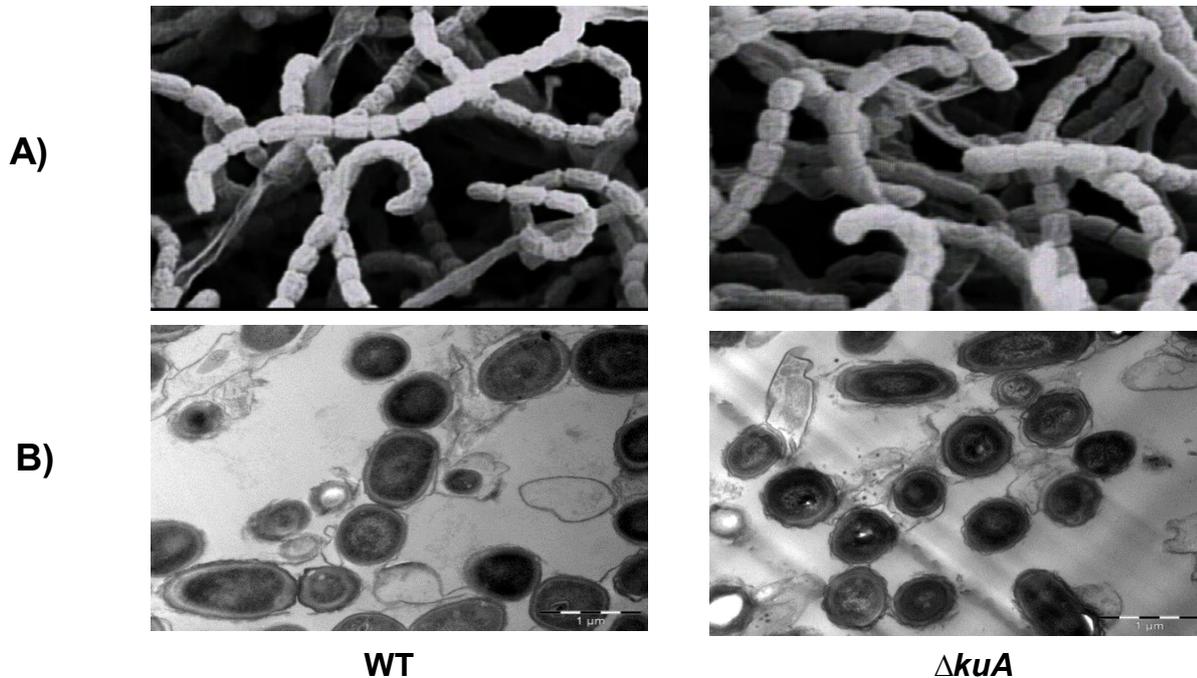


Figure R1 : Observations de chaînes de spores de *Streptomyces* de souches WT et ΔkuA au microscope électronique à balayage (MEB ; **A**) et au microscope électronique à transmission (MET ; **B**)

Exposition aux rayons électroniques

L'exposition aux rayons électroniques des différentes souches mutées pour un gène *ku* (ΔkuA , ΔkuB , et ΔkuC) ou pour les trois ($\Delta kuABC$) a révélé une sensibilité de chacun des trois mutants simples, ainsi qu'une sensibilité accrue pour le triple mutant (Voir article). En parallèle, ce stress génotoxique a aussi été imposé à des doubles mutants $\Delta kuAB$ ($\Delta SAM23877_5082 \Delta SAM23877_6929$) et $\Delta kuAC$ ($\Delta SAM23877_5082 \Delta SAM23877_6942$) (**Figure R2A et B**). Pour chacun des deux contextes génétiques, les résultats reflètent l'exposition de 3 mutants indépendants lors de deux expériences indépendantes. Pour chacun des deux mutants, une sensibilité aux rayons électroniques au moins aussi élevée que celle d'un mutant ΔkuA mais pas plus élevée que celle d'un mutant $\Delta kuABC$ a été observée pour les doses les plus faibles (0.1 à 0.7 kGy).

Les analyses présentées dans l'article ont également permis de constater que le taux de survie aux rayons électroniques diminuait drastiquement dans un mutant $\Delta kuABC \Delta recA$ dépourvu de gènes *ku-like* et muté pour la RH. De la même manière, le graphique de la figure R2C montre qu'un mutant $\Delta ligCD \Delta recA$ dépourvu de ligases NHEJ-like et également muté pour la RH présente une sensibilité semblable au mutant $\Delta kuABC \Delta recA$.

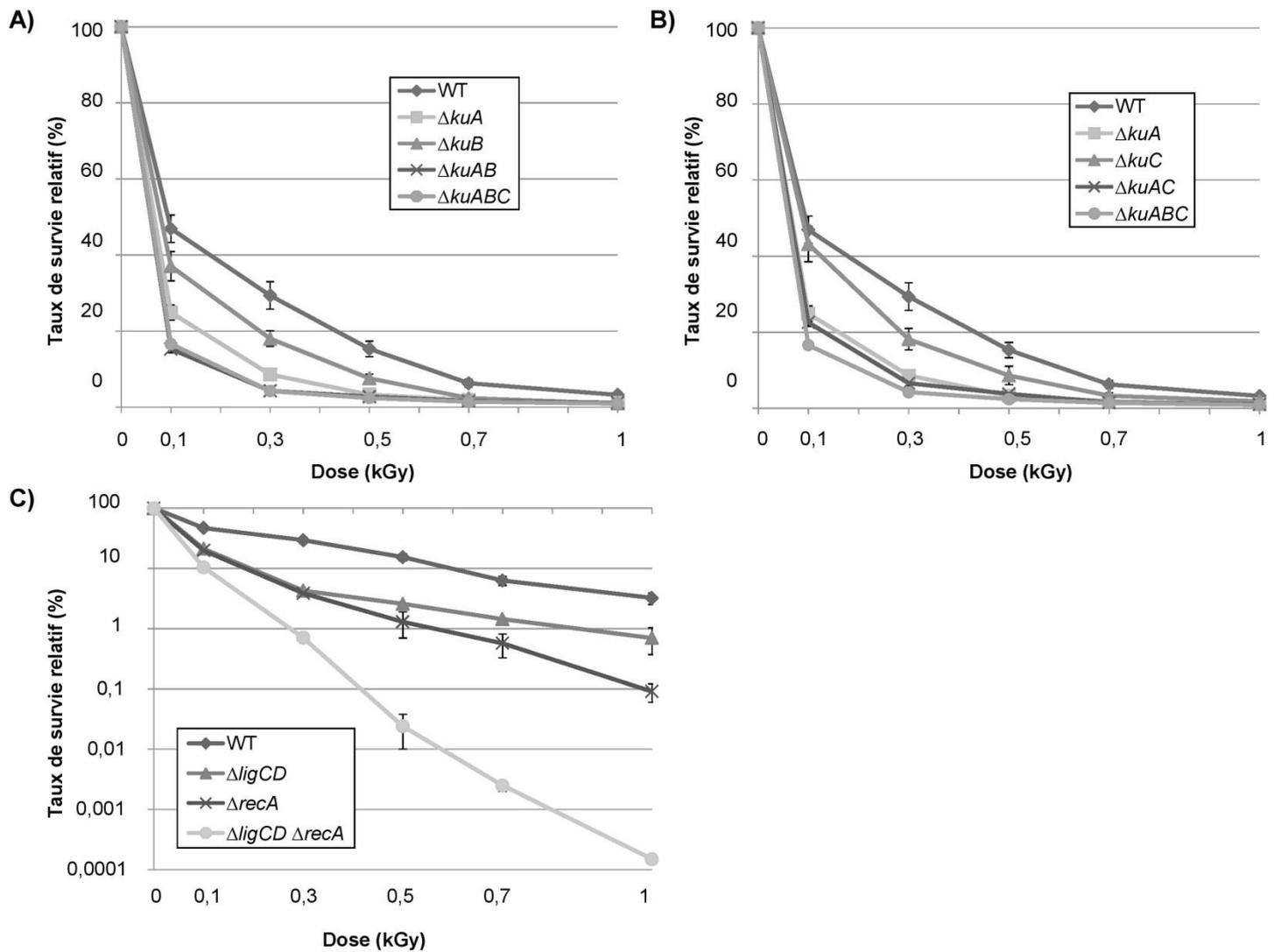


Figure R2: Exposition aux rayons électroniques des souches $\Delta kuAB$ (A), $\Delta kuAC$ (B) et $\Delta ligCD \Delta recA$ (C). Pour chaque graphe, les courbes obtenues pour les simples mutants correspondants sont également représentées, ainsi que pour la souche sauvage qui sert de contrôle. Les spores ont été exposées à 0,1, 0,3, 0,5, 0,7 and 1 kGy de rayons électronique. Les taux de survie relatifs correspondent au rapport du nombre de clones obtenus à une dose donnée sur le nombre de clones non exposés à l'irradiation. La barre d'erreur indique l'erreur standard.

III. Etude des conséquences de la réparation des cassures double-brin sur le génome de *S. ambofaciens*

1. Construction d'un système engendrant des DSB ciblées basé sur l'utilisation de l'endonucléase I-SceI

Afin de tester les conséquences de la génération de DSB le long du chromosome, nous avons développé un système basé sur l'utilisation de la méganucléase I-SceI. Caractérisée initialement chez la levure *S. cerevisiae*, la méganucléase I-SceI catalyse avec une forte spécificité la coupure d'un site non-palindromique de 18 paires de bases (Jacquier et Dujon, 1985). Le système a déjà été utilisé chez de nombreux organismes eucaryotes ou procaryotes, pour l'étude de la réparation des DSB, ou pour d'autres applications biotechnologiques visant à modifier le génome. Du fait de l'absence de sites I-SceI dans le génome des *Streptomyces* et plus généralement des actinomycètes (Siegl *et al.*, 2011 ; Fernandez-Martinez and Bibb, 2014), l'utilisation d'un système I-SceI est un bon outil pour l'étude de la réparation des DSB. En effet, cela permettra d'insérer par construction génétique un site I-SceI unique à l'endroit stratégique souhaité. Toutefois, en raison du faible taux en GC (20%), le gène *sceI* de *S. cerevisiae* ne peut pas être exprimé chez *Streptomyces*. Pour cela, la séquence du gène *sceI* a été optimisée pour une expression chez *Streptomyces* (emploi des codons synonymes majeurs chez *Streptomyces* tout en respectant la séquence primaire de l'endonucléase). Ce gène synthétique, renommé *str-sce*, a été commandé chez un prestataire (GeneCust, Evry, France) et sa séquence est fournie en annexe 1.

Le gène *str-sce* a été cloné dans le plasmide pDYN6902 (Thibessard *et al.*, 2015 ; annexe 2A) sous le contrôle du promoteur inductible au thiostrepton *PtipA*. L'utilisation d'un plasmide à faible nombre de copies tel que pDYN6902 permet de contourner une toxicité de l'expression de *str-sce* chez l'hôte intermédiaire *E. coli*. Le plasmide a, par la suite, été transféré par conjugaison intergénérique chez *S. ambofaciens* où il s'intègre par recombinaison site spécifique au site d'attachement *attB ϕ C31* dans la région centrale du chromosome (*SAM23877_3857* ; position 4.112 Mb ; **figure R3**). L'expression hétérologue du gène *str-sce* a été testée en absence et en présence d'induction au thiostrepton (1 et 2 μ g/mL). Des profils d'expression similaires ont été obtenus avec ou sans induction ce qui suggère que le gène est exprimé même en absence d'inducteur. La fuite de *PtipA* est un phénomène connu qui est lié à une faible répression du promoteur (Chiu *et al.* 1999, Ali *et al.* 2002). Ainsi, dans les expériences réalisées au cours de la thèse avec le système I-SceI, il a

été considéré que le gène *str-sce* est exprimé dès que le plasmide pDYN6902 contenant le gène est introduit dans la souche réceptrice par conjugaison.

Afin de pouvoir générer des DSB ciblées chez *S. ambofaciens*, une cassette servant de cible pour l'enzyme a également été construite. Cette dernière est composée du gène de l'aminoglycoside 3' phosphotransférase (*neo*) conférant la résistance à la kanamycine, flanqué de deux sites *I-SceI*, positionnés de manière convergente afin d'éviter la formation d'extrémités cohésives après coupure par l'endonucléase. Pour l'intégration de la cassette, appelée IKI pour « *I-SceI*/*Kan^R*/*I-SceI* », deux stratégies différentes ont été utilisées :

- Pour générer des DSB dans la région centrale du chromosome, la cassette a été clonée dans le plasmide pMS82 (Grégory *et al.* 2003 ; annexe 2B), qui est conjugatif et à intégration spécifique au site d'attachement du phage ϕ BT1 (*SAM 23877_4663* ; position 4,94 Mb ; **figure R3**). Le locus d'intégration de la région centrale du chromosome sera nommé IKI-C dans la suite du manuscrit.

- Pour générer des DSB dans les bras du chromosome, la cassette a été intégrée au locus cible par la méthode de 'PCR targeting' (Gust *et al.* 2003) dans un espace intergénique (>300 pb) entre deux gènes convergents, afin d'éviter de perturber l'expression de certains gènes. Le locus d'intégration de la cassette dans le bras gauche, situé en position 0,32 Mb, sera nommé IKI-L pour la suite du manuscrit, tandis que le locus d'intégration de la cassette dans le bras droit, situé en position 7,91 Mb, sera nommé IKI-R (**figure R3**). Dans l'ensemble des expériences réalisées, l'étape d'intégration de la cassette au locus choisi est réalisée avant celle de l'intégration du plasmide pDYN6902-*str-sce*.

Afin de tester la fonctionnalité de notre système *in vivo*, un plan expérimental a été mis en place. Ce dernier est basé sur une approche calibrée de conjugaison intergénérique entre des cellules donneuses d'*E. coli* S17 qui contiennent soit le plasmide pDYN6902 vide, soit le plasmide pDYN6902-*str-sce* et des cellules receveuses de *S. ambofaciens* ATCC23877 portant IKI-C. Ainsi, $2 \cdot 10^8$ spores de *Streptomyces* ont été mises en contact avec 10^{10} cellules d'*E. coli* en phase exponentielle de croissance, puis ont été étalées sur milieu SFM après dilution. Après 16 heures de croissance, une pression de sélection a été appliquée pour sélectionner les clones de *S. ambofaciens* ayant intégré le plasmide fourni par *E. coli*. La calibration du nombre de donneuses et de receveuses permet de comparer les fréquences de conjugaison entre elles. Toutefois, étant donné que de nombreux paramètres influent sur la conjugaison, les fréquences de conjugaison se sont révélées très fluctuantes d'une répétition à l'autre. Ainsi il s'est avéré impossible d'établir une moyenne informative des

fréquences de conjugaison pour chacun des deux contextes (présence ou absence de l'enzyme). Néanmoins, au sein d'une même répétition et ce pour chacune d'entre elles (4 répétitions), le nombre de transconjugants obtenus en présence du gène codant l'enzyme *I-SceI* est inférieur à celui obtenu en son absence. Ainsi, lorsque l'on effectue la moyenne du rapport entre le nombre de transconjugants obtenus en présence de l'enzyme et le nombre de transconjugants obtenus en absence de l'enzyme, on constate un facteur 10 ($11,6 \pm 4,6$) entre les deux contextes, ce qui révèle la fonctionnalité de l'enzyme *in vivo*.

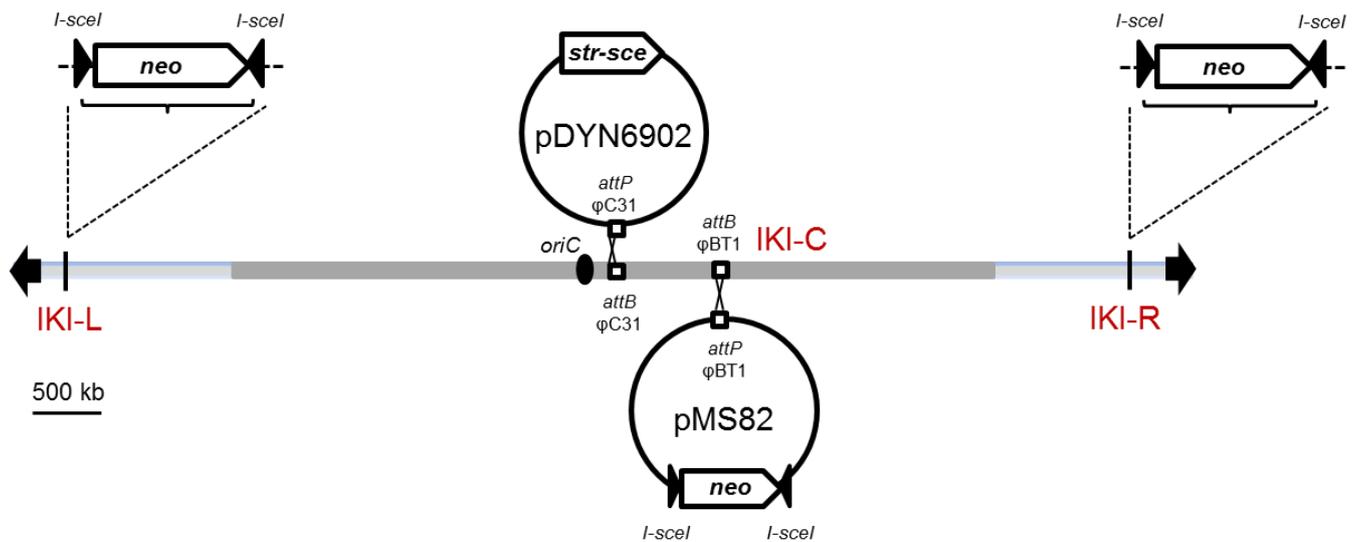


Figure R3 : Localisation des sites d'intégration de la cassette IKI sur le chromosome de *S. ambofaciens*. La cassette IKI, composée du gène *neo* flanqué de deux sites *I-SceI* convergents, est intégrée soit dans la région centrale du chromosome (IKI-C) après clonage dans le plasmide pMS82 qui s'intègre de manière spécifique au site *attB* ϕ BT1, soit dans le bras gauche du chromosome (IKI-L), soit dans le bras droit du chromosome (IKI-R). Le gène codant l'enzyme *str-sce* est porté par le plasmide pDYN6902 qui s'intègre de manière spécifique au site *attB* ϕ C31.

2. La réparation mutagène de DSB dans la partie conservée du chromosome est dépendante du NHEI

La génération de DSB dans la partie centrale du chromosome de *S. ambofaciens* au locus IKI-C a été réalisée dans trois contextes génétiques, à savoir (i) une souche sauvage, (ii) une souche délétée de *kuA*, et (iii) une souche Δ *ligCD*. Ainsi, le mutant Δ *kuA* est dépourvu du seul gène *ku* conservé dans l'ensemble des *Streptomyces*, tandis que le mutant Δ *ligCD* est dépourvu des deux gènes codant des ligases ATP dépendantes NHEJ-like. Dans la partie précédente, il a été démontré que les mutants Δ *kuA* et Δ *ligCD* sont sensibles à l'exposition

aux rayons électroniques, ce qui suggère leur implication dans un mécanisme NHEJ. Après intégration de la cassette IKI (locus IKI-C), l'intégration du plasmide pDYN6902-*str-sce* a permis l'expression de l'enzyme I-SceI. Pour chaque contexte génétique, 50 lignées indépendantes ont été générées en réalisant trois cycles de sous-clonages à partir de transconjugants indépendants. Pour chaque lignée, 200 clones (soit environ 10 000 clones par contexte génétique) ont été testés pour leur sensibilité à la kanamycine. Etant donné que le gène de résistance à la kanamycine est flanqué de deux sites I-SceI, il est attendu que la perte de résistance à la kanamycine résulte soit (i) d'une coupure aux deux sites qui entraîne la perte du gène de résistance suivi d'une réparation illégitime, ou soit (ii) d'une coupure à l'un des deux sites suivi d'une réparation mutagène. Cependant, il est toutefois important de noter que la réparation d'un site I-SceI après coupure n'entraîne pas nécessairement la perte de la résistance à l'antibiotique. Le diagramme de la figure R4 représente le pourcentage de clones sensibles à la kanamycine (Kan^S) pour chaque lignée.

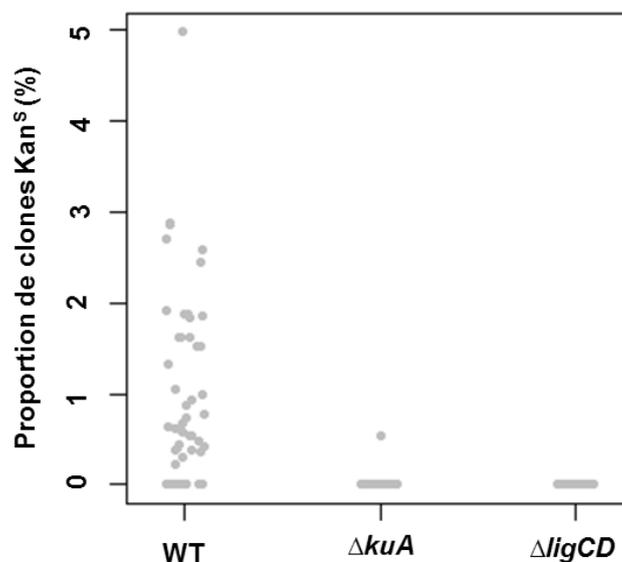


Figure R4 : Diagramme de dispersion représentant la proportion de clones Kan^S pour les lignées WT, ΔkuA et $\Delta ligCD$. Chaque point gris représente une lignée indépendante.

Dans un contexte sauvage, 82 clones Kan^S répartis dans 38 lignées (0 à 10 clones par lignée) ont été obtenus. En revanche, un seul clone Kan^S a été obtenu parmi les lignées délétées pour *kuA* tandis qu'aucun clone Kan^S n'a pu être obtenu dans les lignées $\Delta ligCD$. Les analyses statistiques démontrent que la proportion moyenne de clones Kan^S dans un contexte sauvage (0,99%) est significativement différente des proportions moyennes de clones Kan^S obtenues dans un contexte ΔkuA (0,01%) ou $\Delta ligCD$ (pourcentage nul) ($P < 0.05$). Ainsi, ces résultats (i) démontrent que la perte de résistance à la kanamycine au locus IKI-C

est liée à la présence de *kuA* ainsi que *ligC* et/ou *ligD*, et (ii) suggèrent que ces gènes seraient impliqués dans un mécanisme de NHEJ mutagène.

2.1 L'analyse des clones Kan^S révèle des cicatrices typiques d'une réparation NHEJ

Afin de confirmer l'existence d'un mécanisme de NHEJ responsable d'évènements de réparation mutagène, la séquence nucléotidique de la région entourant la cassette IKI a été précisément analysée dans les clones Kan^S. Pour cela, différents couples d'amorces ont été testés pour obtenir un amplifiat dont le séquençage révélerait la cicatrice de la réparation au locus IKI-C. Etant donné la présence du gène de résistance à l'hygromycine (utilisé pour maintenir la pression de sélection sur le plasmide pMS82), directement en amont de la cassette IKI, les clones ayant subi des délétions dans la partie 5' de la cassette ont été contre-sélectionnés, et seules deux amorces « 5' » situées respectivement 100 et 1000 pb en amont de la cassette ont été nécessaires (**figure R5**). En revanche, aucun gène connu pour être essentiel n'a été détecté directement en aval de la cassette IKI, et des amorces « 3' » ont été dessinées tous les kilobases sur une longueur totale de 20 kb en aval de la cassette (**figure R5**).

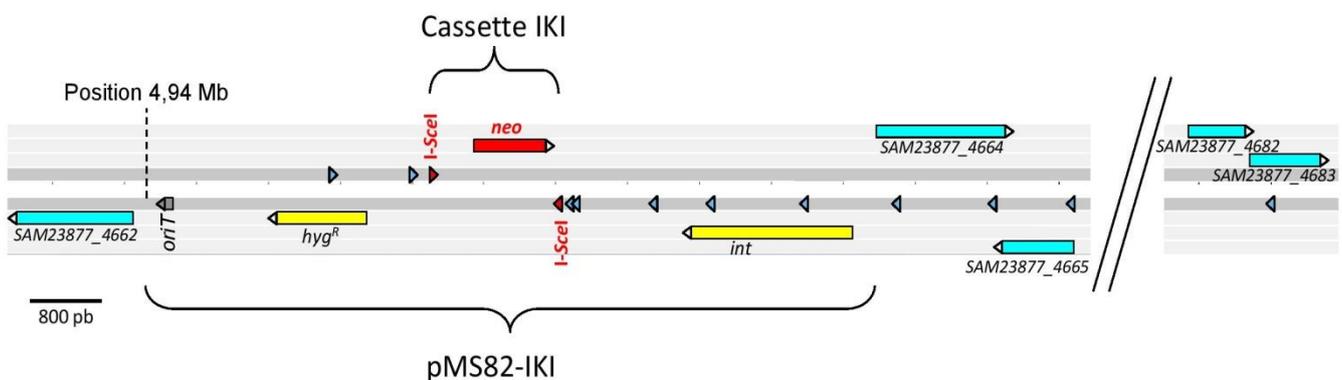


Figure R5 : Environnement génétique du locus de DSB ciblée IKI-C de la région centrale du chromosome. Le plasmide pMS82, intégré au site d'attachement du phage ϕ BT1 en position 4,94 Mb, porte la cassette IKI, composée du gène *neo* qui code la résistance à la kanamycine, flanqué de deux sites I-SceI convergents (en rouge), un gène de résistance à l'hygromycine (*hyg^R*) et le gène *int* codant pour l'intégrase site-spécifique (en jaune), ainsi qu'une origine de transfert par conjugaison (*oriT*). Les ORF du chromosome de *S. ambofaciens* voisines d'IKI-C sont indiquées en bleu. Les amorces utilisées pour séquencer les cicatrices de réparation post-DSB sont représentées par des triangles bleus. Deux amorces « 5' » ont été dessinées en amont de la cassette IKI ainsi qu'un panel d'amorces « 3' » positionnées approximativement tous les kilobases en aval de la cassette sur une longueur totale de 20 kb jusqu'au gène *SAM23877_4683*.

L'amplifiat obtenu pour chaque clone Kan^S a été séquéncé et les cicatrices de réparation ont été analysées en comparant la séquence obtenue avec celle du génome de *S. ambofaciens* ATCC 23877 (Thibessard *et al.*, 2015). Au total, 64 séquences de clones Kan^S ont pu être obtenues pour l'ensemble des lignées du contexte sauvage. Le bilan des caractéristiques des séquences de réparation ainsi que certains exemples représentatifs des cicatrices observées sont présentés respectivement dans les **figures R6** et **R7**.

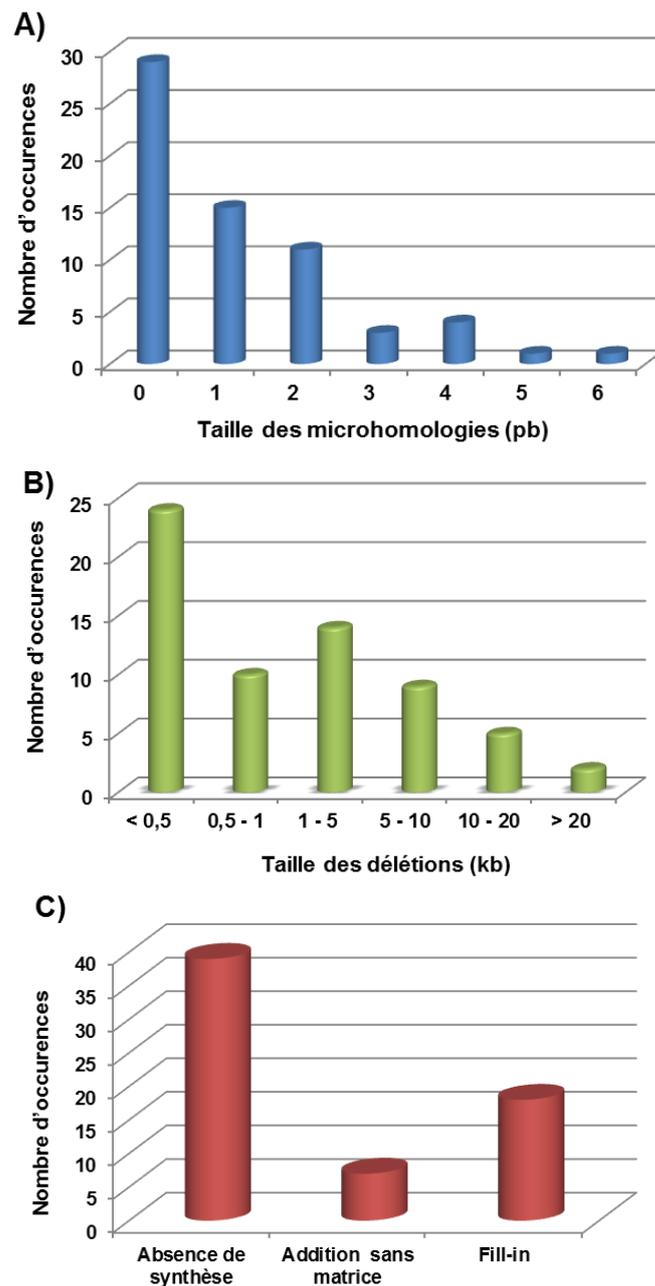


Figure R6 : Caractéristiques des événements de réparation survenus dans les clones Kan^S. A) Taille et occurrence des microhomologies B) Taille et occurrence des délétions C) Nature de l'intervention de la polymérase pour l'ajout de nucléotides.

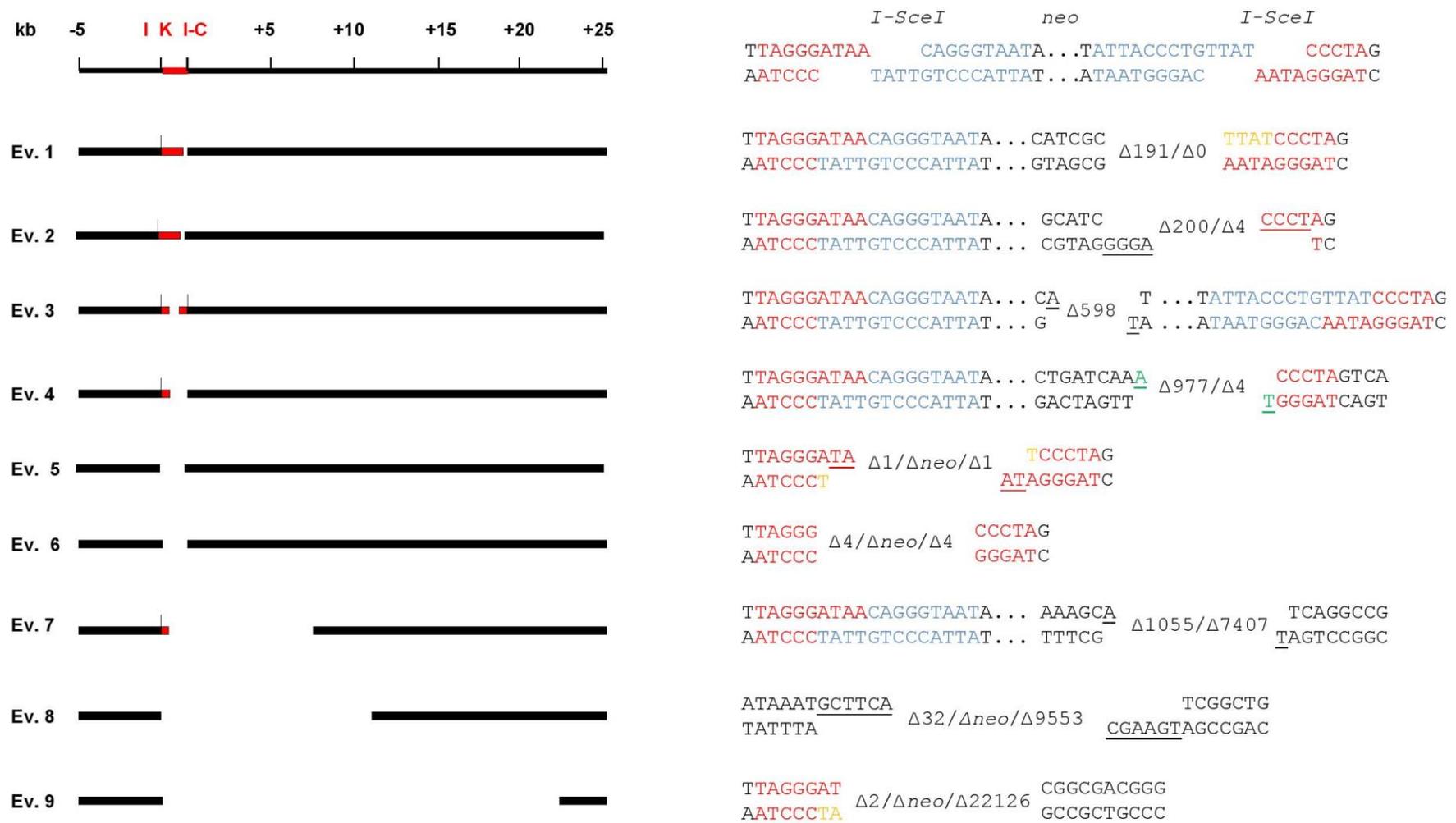


Figure R7 : Exemples caractéristiques de cicatrices de NHEJ obtenues au locus IKI-C. La colonne de gauche représente les délétions (échelle en kb) observées pour les différents évènements de réparation. La cassette IKI est représentée en rouge et les sites I-SceI sont représentés par un trait vertical. La colonne de droite représente la séquence des cicatrices de réparation correspondant aux évènements de la colonne de gauche. La séquence du haut représente la cassette IKI après restriction par la méganucléase I-SceI mais avant réparation. Etant donné que le site I-SceI est non palindromique, les extrémités générées par la restriction sont représentées de deux couleurs différentes (rouge et bleu). Le symbole Δ indique le nombre de nucléotides perdus. Δneo indique une perte du gène de résistance à la kanamycine. Les nucléotides ajoutés en absence de matrice sont représenté en vert, tandis que les nucléotides synthétisés pour combler des lacunes simple-brin (*fill-in*) sont représentés en orange. Les nucléotides intervenant dans la formation de microhomologies sont soulignés.

Pour l'ensemble des séquences analysées, des délétions affectant la cassette IKI ont systématiquement été observées. A l'exception de deux clones pour lesquels la perte de respectivement 180 et 598 nucléotides au sein du gène de résistance à la kanamycine semble indépendante de la restriction d'I-SceI (les deux sites sont intacts, **figure R7**, ev. 3), tous les évènements de réparation incluent une modification d'un site (25/62) ou des deux (37/62) sites I-SceI. Dans la majorité des cas, la résection des extrémités de la cassure se fait de manière bidirectionnelle c'est-à-dire que les deux extrémités de la cassure subissent une délétion (55/62) et plus rarement, la résection des extrémités se fait uniquement de manière unidirectionnelle (**figure R7**, ev. 1). La taille des délétions est variable : pour plus d'un tiers (24/62), moins de 500 pb sont perdues (191 pb pour la délétion minimale observée ; **figure R7**, ev. 1), tandis que pour 7 clones sur 62, plus de 10 kb ont été perdus (23,6 kb pour la délétion la plus importante observée, soit 20 gènes ; **figure R7** ev. 9). Pour 7 clones Kan^S, aucune amplification PCR n'a été obtenue, suggérant que la délétion excède 23,6 kb.

Pour un peu plus de la moitié des cas (35/64), la cicatrization de la cassure a été favorisée par la présence de microhomologies résultant du traitement des extrémités. Dans la majorité de ces 35 cas, la taille des microhomologies varie entre 1 et 3 nucléotides (29/35 ; **figure R7**, ev. 7). Cependant, l'analyse de certaines séquences a révélé l'utilisation de microhomologies dont la taille varie entre 4 et 6 nucléotides (6/35 ; **figure R7**, ev. 2 et 8). Contrairement aux évènements de délétion, observés dans 100% des cas, l'addition de nucléotides a été observée dans moins de la moitié des cas (26/64) Ceux-ci peuvent être classés en deux catégories. Dans la première catégorie (8/26), un seul nucléotide a été ajouté au niveau d'extrémités franches formées par une résection bidirectionnelle (**figure R7**, ev. 4). Dans la seconde catégorie (18/26), l'addition de 1 à 4 nucléotides se fait par « *fill-in* », c'est-à-dire par comblement de lacunes simple-brin générées lors de la réparation de la cassure (**figure R7**, ev. 1, 5 et 9).

Il est intéressant de noter que certaines cicatrices de réparation ont été observées à plusieurs reprises. C'est notamment le cas des évènements 5 et 6 de la **figure R7**, observés respectivement 3 et 2 fois. Dans les deux cas, les deux sites I-SceI ont été coupés et la cassette IKI a été perdue. Dans le cas de l'évènement 5, un seul nucléotide a été délété des extrémités 3' sortantes, formant ainsi une microhomologie de 2 nucléotides. Avant la ligation, la lacune simple-brin d'un nucléotide a été comblée par *fill-in*. Pour l'évènement 6, les 4 nucléotides des extrémités 3' sortantes ont été délétés et les extrémités franches ainsi engendrées ont été liguées. La ligation d'extrémités incompatibles après traitement

(résection, resynthèse) est une des caractéristiques du mécanisme de NHEJ tel qu'il a été caractérisé chez *M. tuberculosis* (Stephanou *et al.*, 2007 ; Aniuoku *et al.*, 2008 ; Gupta *et al.*, 2011).

2.2 Le NHEJ permet l'insertion de matériel génétique ectopique au site de cassure

Pour trois clones particuliers, la jonction d'extrémités lors de la réparation de la cassure est accompagnée de l'intégration de courts fragments d'ADN (30, 42 et 83 nucléotides ; **figure R8**).

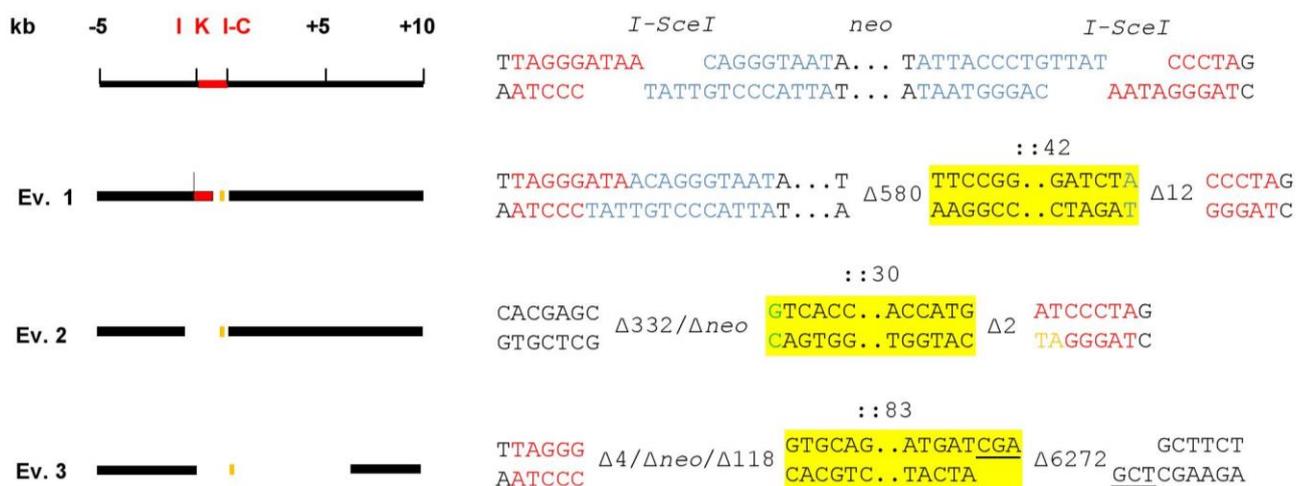


Figure R8 : Représentation des cicatrices de réparation NHEJ obtenues au locus IKI-C et ayant permis l'intégration de fragments d'ADN au site de cassure. La colonne de gauche représente les délétions (échelle en kb) observées pour les différents évènements de réparation. La cassette IKI est représentée en rouge, les sites I-SceI sont représentés par un trait vertical, et les évènements d'insertion sont représentés en jaune. La colonne de droite représente la séquence des cicatrices de réparation correspondante aux évènements de la colonne de gauche. La séquence du haut représente la cassette IKI après restriction par la méganucléase I-SceI mais avant réparation. Etant donné que le site I-SceI est non palindromique, les extrémités générées après restriction sont représentées de deux couleurs différentes (rouge et bleu). Le symbole Δ indique le nombre de nucléotides perdus et les :: indiquent le nombre de nucléotides insérés. Δ_{neo} indique une perte du gène de résistance à la kanamycine. Les nucléotides ajoutés en absence de matrice sont représentés en vert, tandis que les nucléotides synthétisés pour combler des lacunes simple-brin (*fill-in*) sont représentés en orange. Les nucléotides intervenant dans la formation de microhomologies sont soulignés.

Par exemple, pour l'évènement 3 (**figure R8**), la cassette IKI a été perdue et l'extrémité 3' du site de gauche a été délétée. L'insertion du fragment de 83 nucléotides pourrait être facilitée par une microhomologie de 3 nucléotides en 3'. Les séquences de 42 et 83 nucléotides s'alignent parfaitement dans la région centrale du chromosome de *S. ambofaciens*, localisée à proximité du site d'intégration d'IKI-C et délétée durant la réparation. En revanche, pour le fragment de 30 nucléotides (ev. 2, **figure R8**), la séquence s'aligne parfaitement avec la partie 5' d'un gène codant une transposase localisé dans les répétitions terminales du chromosome. Ces résultats suggèrent que le NHEJ des *Streptomyces* est un mécanisme pouvant contribuer à l'insertion de matériel génétique.

2.3 *kuA* serait un acteur majeur mais non exclusif du NHEJ

Lors du criblage visant à déterminer la résistance ou la sensibilité à la kanamycine des différents clones, un seul clone Kan^S a pu être identifié parmi les lignées ΔkuA . L'analyse de sa séquence au point d'intégration de la cassette a révélé une délétion bidirectionnelle de 339 nucléotides affectant un des deux sites *I-SceI*, et une ligation survenue au niveau d'une microhomologie de 2 nucléotides. Ces résultats montrent que même si la fréquence de réparation mutagène par NHEJ diminue fortement en absence de *kuA*, la formation de cicatrices de réparation typiques du NHEJ est toujours possible. Il est envisageable que l'absence de KuA puisse partiellement être compensée par KuB et/ou KuC.

2.4 L'analyse des clones Kan^R confirme les résultats observés pour les clones Kan^S

Etant donné la position des sites *I-SceI* par rapport au gène *neo* (485 pb en amont de l'ATG pour l'un, 148 pb en aval du codon stop pour l'autre), la réparation mutagène de l'un ou des deux sites n'entraîne pas nécessairement une inactivation du gène de résistance à la kanamycine. Ainsi, le locus IKI-C d'un ou deux clones Kan^R a été analysé pour chaque lignée. Pour l'ensemble des clones Kan^R analysés dans les contextes ΔkuA et $\Delta ligCD$, les sites *I-SceI* se sont révélés intacts, suggérant une absence de réparation mutagène. En revanche, sur les 92 clones testés issus de lignées du contexte sauvage, 12 présentent un (11/12) ou deux (1/12) sites *I-SceI* modifiés après réparation par NHEJ. L'ensemble de ces 13 sites présentent des délétions inférieures à 100 pb qui n'affectent pas la séquence codante du gène *neo*. Comme pour les clones Kan^S, le comblement de lacunes simple-brin (5/13) ou l'addition sans matrice d'un nucléotide aux extrémités de la cassure (2/13) a également été observé. Dans plus de la moitié des cas (7/13), la ligation pourrait également être médiée par la présence de microhomologies. Ces résultats montrent que les proportions des

différents évènements moléculaires ne sont pas différentes entre les deux catégories phénotypiques (Kan^S et Kan^R).

3. La réparation des DSB dans les bras chromosomiques induit de l'instabilité génétique et génomique

3.1 Analyses des phénotypes coloniaux

Comme pour la région centrale du chromosome, l'expérience de génération de cassures ciblées dans les bras chromosomiques aux loci IKI-L et IKI-R a été réalisée dans des contextes WT, *ΔkuA* et *ΔligCD*. Pour chaque bras chromosomique et pour chaque contexte génétique, 20 lignées indépendantes ont été générées afin de pouvoir analyser les conséquences de réparation des DSB en fonction de leur localisation et de la présence ou non de certains gènes clés du NHEJ. En parallèle, 30 lignées témoins (10 pour chaque contexte génétique) comportant uniquement l'enzyme ont également été réalisées. Après 3 sous-clonages successifs, quel que soit le contexte génétique étudié, il s'est avéré que la plupart des lignées présentait une proportion majeure de clones Kan^S ce qui suggère que le locus IKI-L ou IKI-R a été affecté et que le gène de résistance à la kanamycine a été muté. Si certaines lignées montrent une population au phénotype homogène (**figure R9A**), d'autres en revanche présentent une population avec des phénotypes coloniaux hétérogènes (**figure R9B**), suggérant un mélange de génotypes résultant probablement de différents évènements de réparation. Ainsi le nombre de sous-clonages a été adapté (3 à 8 en fonction des lignées) afin d'homogénéiser les phénotypes.

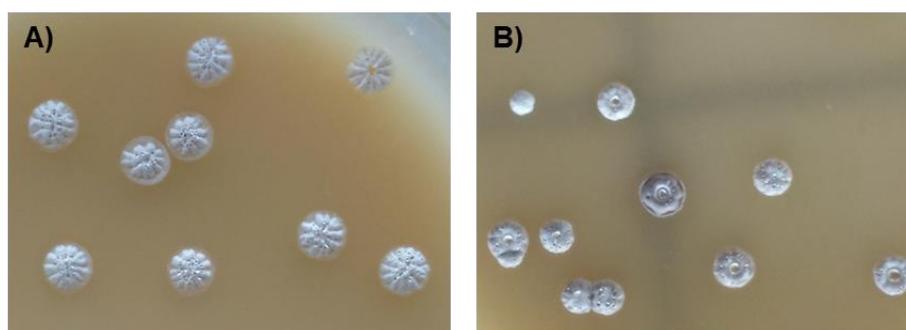


Figure R9 : Phénotypes coloniaux de lignées après génération de DSB dans un bras chromosomique (A) lignée au phénotype colonial homogène (B) lignée au phénotype colonial hétérogène.

En revanche, d'une lignée à l'autre, les phénotypes coloniaux observés sont variables (**figure R10**). En effet, la morphologie des colonies, leur taille, la pigmentation du mycélium

ainsi que la sporulation sont les principales caractéristiques variables observées. Ainsi, ces résultats contrastent avec les expériences effectuées dans la région centrale du chromosome.

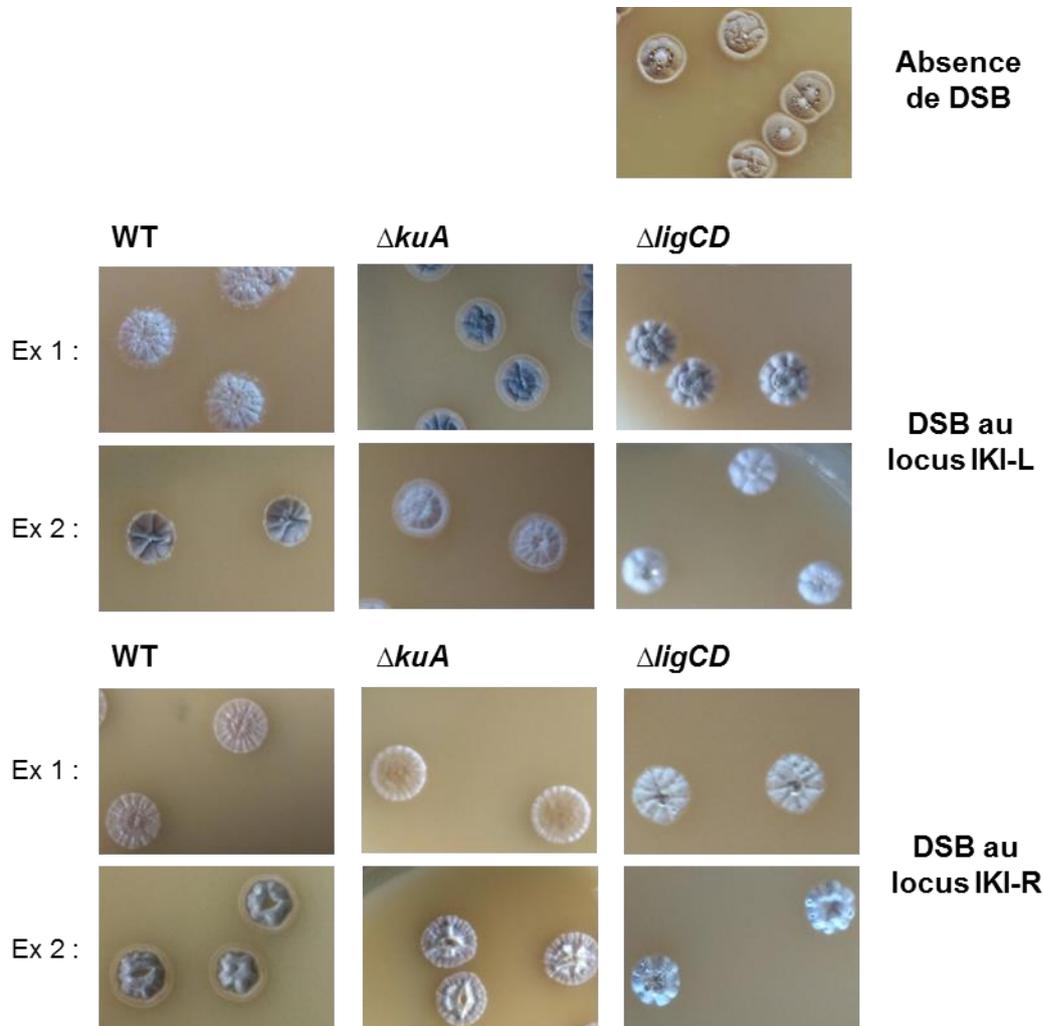
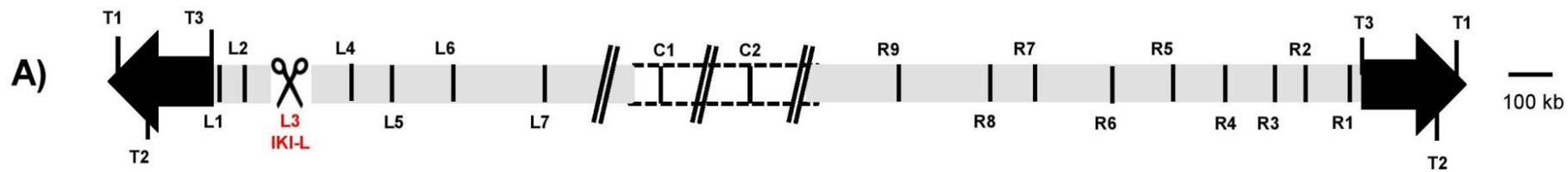


Figure R10 : Variabilité des phénotypes coloniaux observés entre différentes lignées après génération de DSB dans les bras chromosomiques.

3.2 Observation/analyse de l'intégrité des génomes par PCR

Afin de tester l'intégrité des génomes des différentes lignées, une approche préliminaire basée sur l'amplification de loci répartis le long du chromosome a été mise en place. La localisation des différents couples d'amorces est présentée dans la **figure R11**. L'analyse des génomes des différentes lignées révèle des profils variables, avec l'absence de certains loci, suggérant des délétions, et ce quel que soit le contexte génétique. Pour l'ensemble des lignées testées (150), les loci localisés dans la partie centrale du chromosome (C1 et C2) sont toujours présents.



bras gauche						centre						bras droit												
T1	T2	T3	L1	L2	L3	L4	L5	L6	L7	C1	C2	R9	R8	R7	R6	R5	R4	R3	R2	R1	T3	T2	T1	x/60
+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	3	
+	+	+	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	1	
+	+	+	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	11	
+	+	+	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	1	
-	-	+	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	1	
-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	1	
-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	1	
-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	11	
-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-	4	
-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	-	-	-	-	1	
-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-	1	
-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-	5	
-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-	1	
-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-	1	
-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-	12	
-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	-	-	-	-	1	
-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-	2	
-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	-	-	-	-	1	

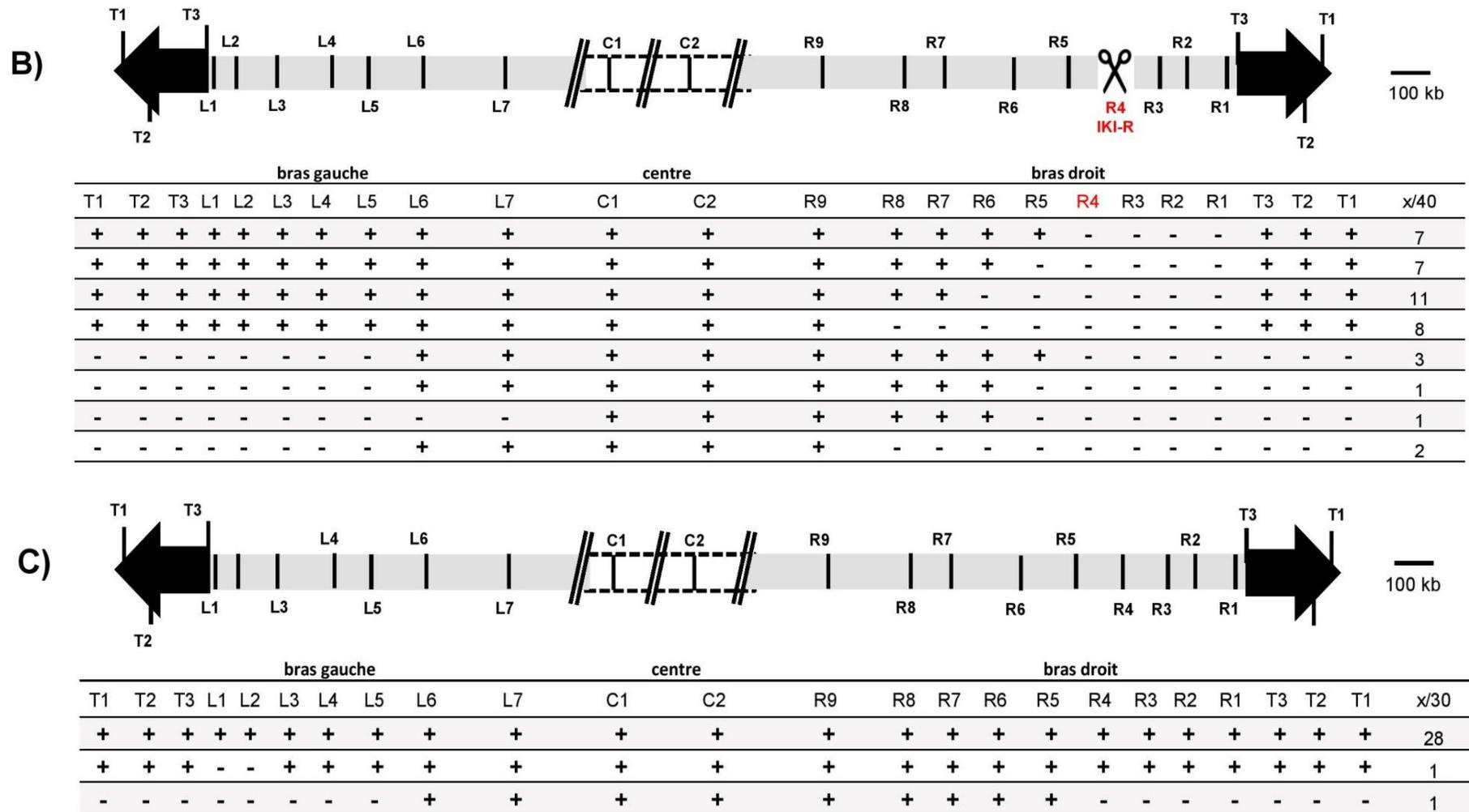


Figure R11: Analyse de l'intégrité des génomes de lignées dans lesquelles est exprimé le gène *str-sce* en présence de la cassette IKI intégrée dans le bras gauche du chromosome (A), dans le bras droit du chromosome (B), ou en absence de la cassette (C). Sur le schéma du chromosome, les bras du chromosome apparaissent en gris, la région centrale en blanc encadrée de pointillés, et les répétitions terminales (TIR) sont représentées par des flèches noires. Les doubles barres indiquent une discontinuité dans la représentation du chromosome. Le site de cassure est représenté par des ciseaux et est indiqué en rouge. Les traits verticaux indiquent les positions des différents couples d'amorces testés ; T indique une localisation dans les TIR, L dans le bras gauche, C dans la région centrale et R dans le bras droit. Dans les tableaux, les + et les - indiquent respectivement la présence ou l'absence d'un signal pour les couples d'amorces correspondant. La colonne de droite indique le nombre de fois où les différents profils ont été observés.

Si l'on s'intéresse de plus près aux 60 lignées où la DSB a été générée dans le bras gauche (**figure R11A**), à l'exception de trois d'entre elles où un signal a été détecté pour l'ensemble des loci, aucun signal n'a été obtenu pour le locus L3 bordant la cassette IKI. Ces résultats confirment ainsi l'observation systématique de sensibilité à la kanamycine et suggèrent une perte du gène *neo*. Par ailleurs les loci L1 et L2 localisés entre la cassette IKI et l'extrémité du bras gauche ainsi que L4, premier locus en aval du site IKI-L sont systématiquement absents (à l'exception de deux lignées pour L4), suggérant des délétions liées à des remaniements de grande ampleur. Pour certaines lignées, le locus L5 et plus rarement le locus L6, sont également absents. Lorsque les loci localisés dans les TIR sont détectés (T1 à T3), il n'est pas possible de conclure sur la présence ou l'absence de l'extrémité du bras gauche, étant donné que le bras droit portant une copie des séquences formant la TIR donne un signal pour ces loci. Dans ces cas-là, la linéarité du chromosome est probablement maintenue. Dans 44 lignées sur les 60, aucun signal n'a été obtenu pour les marqueurs terminaux des deux bras du chromosome (T1-T3), ainsi que pour certains loci du bras droit (R1 à R6 en fonction des lignées), révélant une perte probable d'une grande partie des deux bras du chromosome et une circularisation du chromosome.

Après génération d'une DSB au locus IKI-R, les analyses PCR révèlent pour les lignées des contextes WT et ΔkuA des résultats similaires à ceux obtenus dans le bras gauche, à savoir des profils d'amplification présentant des délétions affectant IKI-R (R4) ainsi que les loci situés en aval (R1-R3), et en fonction des lignées certains loci localisés entre IKI-R et la région centrale du chromosome (R5 à R8) (**figure R11B**). Toutefois, seules 7 lignées sur 40 présentent des profils d'amplification suggérant une circularisation du chromosome, avec absence des loci terminaux (T1 à T3) et de certains loci du bras gauche (L1 à L7, en fonction des lignées). De manière surprenante, les résultats obtenus dans les lignées $\Delta ligCD$ sont totalement différents avec 19 lignées sur 20 présentant une circularisation probable du chromosome, et pour lesquelles le dernier locus amplifié sur le bras gauche est systématiquement L4. Le profil d'amplification de la 20^{ème} lignée est encore plus surprenant puisqu'il révèle un bras droit intact et uniquement une absence des loci L1 à L4. Cela nous a amené à penser que la lignée ancestrale utilisée pour intégrer la cassette IKI a été remaniée préalablement à l'intégration des constructions. La vérification de l'intégrité du génome de cette lignée a révélé un profil d'amplification avec absence des loci L1 à L4, confirmant un réarrangement antérieur à l'intégration du système I-SceI. Pour cette raison, les résultats obtenus pour ce contexte génétique n'ont pas été intégrés pour l'analyse des profils dans la **figure R11B**. Actuellement, 20 nouvelles lignées issues d'une souche $\Delta ligCD$ dont l'intégrité

chromosomique a été vérifiée sont en cours de construction au laboratoire, afin de pouvoir compléter les analyses de la réparation de DSB dans le bras droit du chromosome.

L'analyse PCR des lignées témoins (où l'enzyme est exprimée en absence de la cassette) a révélé la présence de la totalité des loci testés pour 28 des 30 lignées (**figure R11C**). En effet, l'absence des loci L1 et L2 suggère un remaniement impactant l'extrémité du bras gauche du chromosome pour une première lignée, tandis que l'absence de signal pour l'ensemble des loci en amont de L6 sur le bras gauche et en aval de R5 sur le bras droit suggère une circularisation du chromosome accompagnée de larges délétions pour une seconde lignée. Ces réarrangements traduisent la forte fréquence de réarrangements spontanés du chromosome de *S. ambofaciens*.

3.3 Analyse des évènements de réparation

Afin de déterminer précisément la structure du génome et la nature des évènements accompagnant la réparation de DSB, le génome de 21 lignées (7 par contexte génétique) a été séquencé. Etant donné que la réparation de DSB entraîne des conséquences similaires lorsqu'elle est générée au locus IKI-R ou au locus IKI-L, nous avons choisi de nous focaliser uniquement sur des génomes où la cassure est localisée au locus IKI-L. Le séquençage a été réalisé grâce à la technologie Illumina par la méthode dite de « *paired-ends* ». La préparation de banques de séquençage pour cette méthode consiste à fragmenter le génome pour atteindre des tailles 300 à 400 pb. L'ensemble des *reads* obtenus ont par la suite été alignés avec le génome de référence de la souche ATCC 23877 (Thibessard *et al.* 2015). Ainsi, la taille des régions délétées ou dupliquées, ainsi que les bornes de ces dernières ont pu être identifiées dans la majorité des cas. Afin de faciliter la compréhension des évènements décrits dans les paragraphes qui suivent, la **figure R12** présente la localisation des différents loci impliqués.

3.3.1 Réparation de DSB par remplacement de bras chromosomique

Pour 6 des 21 souches, l'analyse des séquences révèle une délétion d'une partie du bras gauche qui se traduit par la chute du taux de couverture des *reads* entre les positions 481 000 et 210 000 (qui correspond au début des TIR). L'observation de *reads* entre la position 210 000 et l'extrémité du bras gauche pourrait résulter de la présence d'une copie intacte des TIR sur le bras droit. Ainsi, cela suggérerait une perte de l'extrémité du bras gauche située en aval de la position 481 000. A l'inverse, il a été constaté que le taux de couverture des *reads* double à partir de la position 7 474 400 (bras droit) jusqu'à l'extrémité

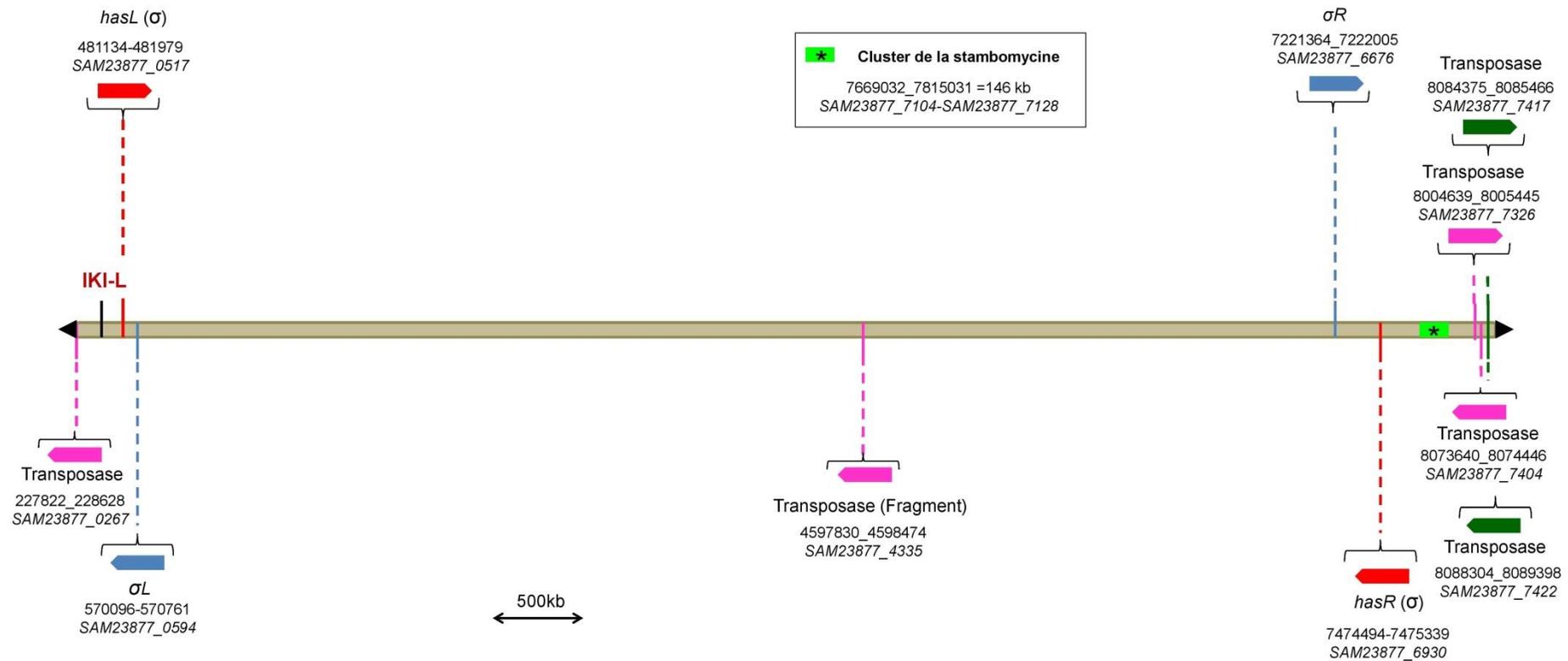


Figure R12 : Représentation des différents loci impliqués dans des évènements de RH après génération d'une cassure au site IKI-L. Pour chaque locus, l'orientation, la position sur le chromosome et le locus-tag sont indiqués. Les triangles noir aux extrémités du chromosome indiquent la position des TIR.

du chromosome (**figure R13**). Les positions 481 000 et 7 474 400 correspondent aux gènes codant respectivement les facteurs σ HasL (*SAM023877_0517*) et HasR (*SAM23877_6930*) qui partagent 99% d'identité nucléotidique (823/835 nt). Par le passé, il a été démontré que ces loci, qui sont positionnés en orientation opposée, pouvaient être impliqués dans des événements de remplacement de bras initié par un mécanisme de type *Break-Induced Replication* (BIR) (Fischer *et al.* 1998).

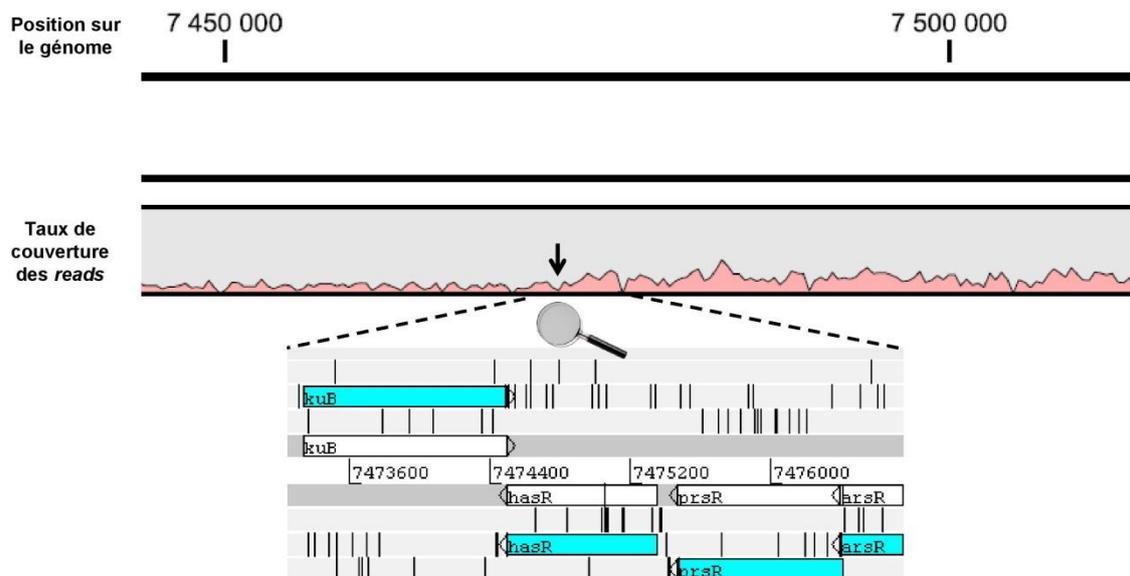


Figure R13 : Augmentation du taux de couverture des *reads* au niveau du gène *hasR* caractérisant un événement de duplication du bras.

Ainsi, l'hypothèse avancée pour ces 6 lignées est que le bras droit aurait remplacé le bras gauche délété, après un événement de type BIR entre les gènes *has* dupliqués. Ainsi, la perte des 481 derniers kb du bras gauche serait compensée par la duplication des 830 derniers kb du bras droit. La taille totale du génome passerait alors de 8,303 kb pour la souche sauvage à 8,648 kb dans ces souches réarrangées. Les trois étapes hypothétiques conduisant au remplacement du bras gauche par le bras droit du chromosome de *S. ambifaciens* grâce au mécanisme de BIR sont illustrées en **figure R14A**.

3.3.2 Circularisation du génome

Pour les 15 autres génomes parmi les 21, les deux bras du chromosome présentent des délétions variant de 683 kb à 2 116 kb, et affectant les deux extrémités du chromosome, suggérant une perte de la linéarité. La conformation circulaire du chromosome a été confirmée pour la majorité de ces souches après identification des points de recombinaison.

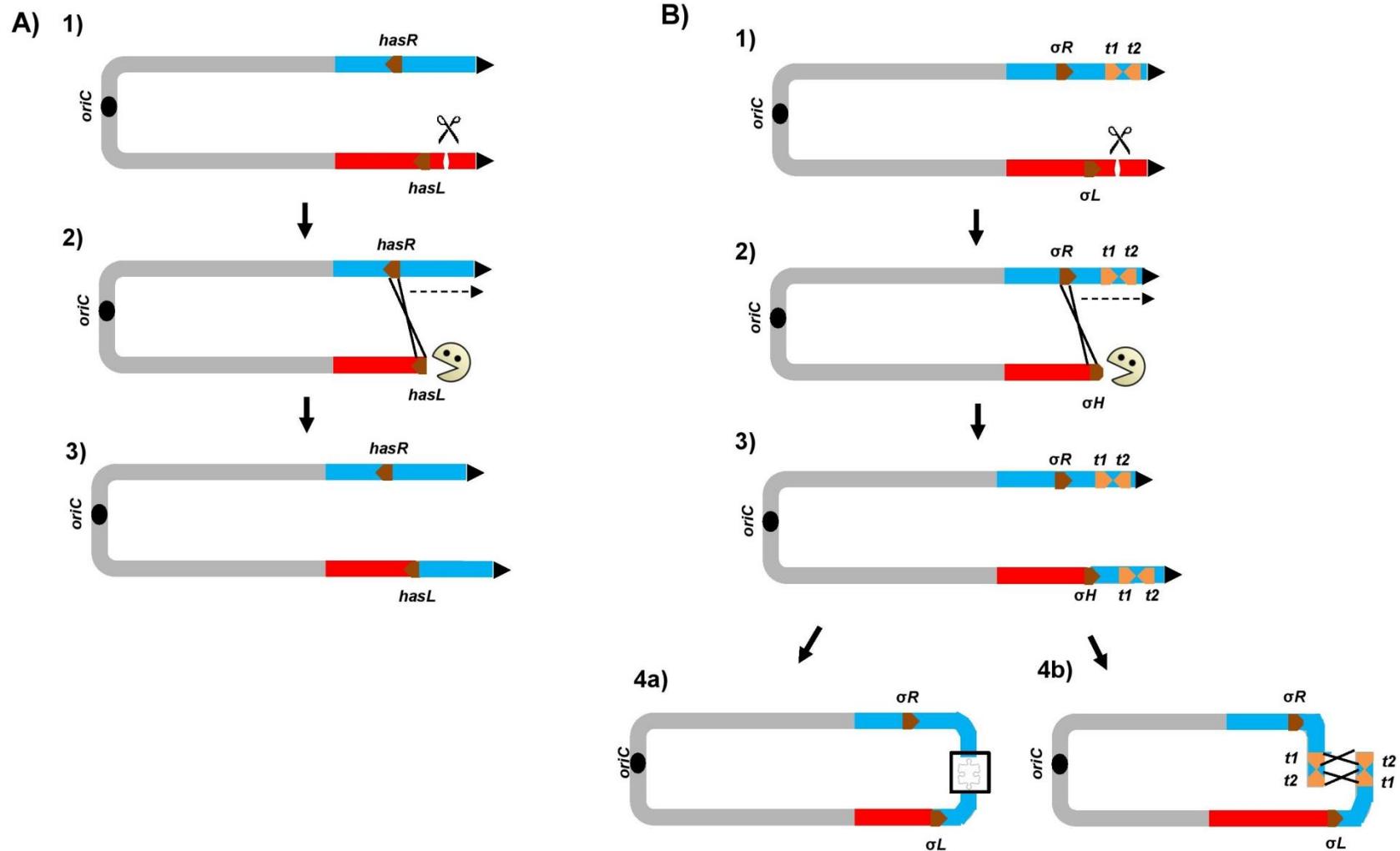


Figure R14 : Evènements de recombinaison simples ou multiples entraînant des réarrangements génomiques majeurs après génération d'une DSB au locus *IKI-L*. **A)** Après génération de la cassure (ciseaux) (1), les extrémités sont traitées (« *pacman* ») et un évènement de BIR entre les deux copies du gène *has* (marron) présent sur chaque bras chromosomique permet le remplacement du bras gauche endommagé (rouge) par une duplication du bras droit intact (bleu) (2-3). **B)** Les étapes 1 à 3 initiales sont identiques à celles présentées en A), mais le gène utilisé comme matrice pour le BIR est différent (σ_L et σ_R , marron). Toutefois, la configuration linéaire n'est jamais maintenue, et la circularisation du chromosome peut intervenir soit par un évènement de fusion illégitime entre deux extrémités traitées des deux copies du bras droit (4a), soit par recombinaison homologue entre le gène de transposase *t1* et sa copie homologue *t2* localisée sur la duplication du bras droit ou inversement.

Il est intéressant de noter que pour 11 de ces souches, la délétion sur le bras droit s'achève en position 570 000, pour laquelle l'analyse du génome révèle la présence d'un gène codant un facteur σ que nous avons nommé σL (*SAM023877_0594*). En criblant le génome de la souche de référence, il a été constaté qu'une copie dupliquée de ce gène est présente en orientation opposée en position 7 220 000 sur le bras droit du chromosome (nommé σR , *SAM23877_6676*). La région dupliquée est chevauchante avec la séquence codante adjacente, et constitue une séquence répétée de 1673 pb avec 99% d'identité nucléotidique. En analysant le taux de couverture des *reads* du génome, il a été constaté que ce dernier double à partir de ce locus et ce jusqu'au point de circularisation, suggérant fortement qu'un évènement de BIR se soit déroulé avant la circularisation du chromosome (**figure R14B**). Ainsi, la modification du chromosome résulterait de réarrangements multiples, avec dans un premier temps un évènement de BIR impliquant les gènes dupliqués σL and σR , et dans un deuxième temps une circularisation du chromosome. Dans 3 cas sur 11, des séquences de fusion dont les extrémités sont présentes en orientation opposée sur le bras droit ont pu être identifiées, suggérant que ces évènements de réparation par RI seraient uniquement possibles après duplication du bras droit par un évènement de BIR (**figure R14B-4a**). Dans 3 autres cas sur 11, l'analyse des génomes révèle que la circularisation du chromosome serait liée à des évènements de RH entre des gènes de transposase dupliqués. De la même manière que pour la RI, ces paires de loci (*SAM23877_7417/SAM23877_7422* ou *SAM23877_7326/SAM023877_7404* ; représentés par *t1* et *t2* sur la **figure R14B-4b** ; **figure R12**) sont positionnés sur le bras droit en orientation opposée, et la circularisation par RH est uniquement possible s'il y a préalablement eu un évènement de BIR entre σL et σR qui entraîne une duplication du bras droit (**figure R14B-4b**). Pour 1 cas sur 11, une séquence de 41 nucléotides présentant 95% d'identité nucléotidique a pu être identifiée aux deux extrémités de la fusion. Etant donné que la taille minimale d'homologie requise pour la survenue d'un évènement de RH n'est pas connue chez les *Streptomyces*, il est difficile de conclure si la circularisation découle d'un évènement de RH ou de RI. Pour les 4 derniers cas sur 11, la nature de l'évènement de circularisation reste indéterminée malgré une analyse minutieuse des bornes de la délétion.

Pour les 4 dernières lignées sur les 15 présentant un chromosome circulaire, aucun évènement initial de BIR n'a pu être détecté. Dans les deux premiers cas, une séquence de fusion entre l'extrémité non délétée du bras gauche et l'extrémité non délétée du bras droit a pu être déterminée, révélant ainsi un évènement de RI. Il est intéressant de noter que pour une de ces deux souches, 2 116 kb d'ADN ont été perdus, ce qui correspond à plus de 25%

du génome. Enfin, pour les deux dernières lignées, au moins un évènement de RH et un évènement de RI ont pu être décelés, mais la conformation du génome reste indéterminée.

3.3.3 L'amplification d'ADN dans le cluster de gènes de la stambomycine comme conséquence de la réparation des DSB

Dans 5 génomes sur les 21, une forte augmentation du taux de couverture des *reads* a été observée dans une petite partie du bras droit du chromosome, suggérant que le locus est présent en grand nombre de copies. Lorsque l'on s'intéresse de plus près à la localisation de ces différents loci, il a été constaté que bien qu'ils soient de taille variable (de 14 à 90 kb), ils sont systématiquement localisés dans des gènes de polyketides synthétases (PKS) du cluster de synthèse des stambomycines. Les PKS contiennent des domaines enzymatiques organisés en modules, et la présence de domaines fonctionnels identiques dans les modules successifs du cluster fournit de nombreuses répétitions de motifs nucléotidiques hautement homologues.

Du fait de la nature répétée des séquences du cluster, il est envisageable que la présence de multiples copies des différents loci s'apparente à de l'amplification d'ADN, un phénomène connu chez les *Streptomyces* qui entraîne la réitération en tandems directs de motifs amplifiables appelés AUD (unité amplifiable d'ADN ; Leblond *et al.* 1990). La localisation et les caractéristiques des potentielles amplifications observées sont présentées dans la **figure R15** et dans le **tableau R1**. Parmi les 5 exemples, 3 AUD de 14,2 kb se sont avérées identiques (ex. 1, 3 et 5). Après analyse des bornes de l'amplification, nous avons pu détecter à chaque extrémité une séquence répétée de 1 167 pb (95% d'homologie) positionnée en orientation directe et donc susceptible d'être à l'origine de l'amplification selon le modèle proposé par Young et Cullum (1985). Pour les exemples 2 et 4, la taille de l'AUD est respectivement de 89,7 kb et 32,8 kb. Deux séquences répétées positionnées en orientation directe et susceptibles d'être impliquées dans l'amplification (723 pb, 91% d'homologie pour l'ex. 2, 740 pb, 96% d'homologie pour l'ex. 5) ont été retrouvées aux bornes des AUD respectives. Il est important de noter que pour les 5 exemples, les bornes d'un évènement de circularisation dont la nature reste inconnue sont localisées au sein de l'AUD.

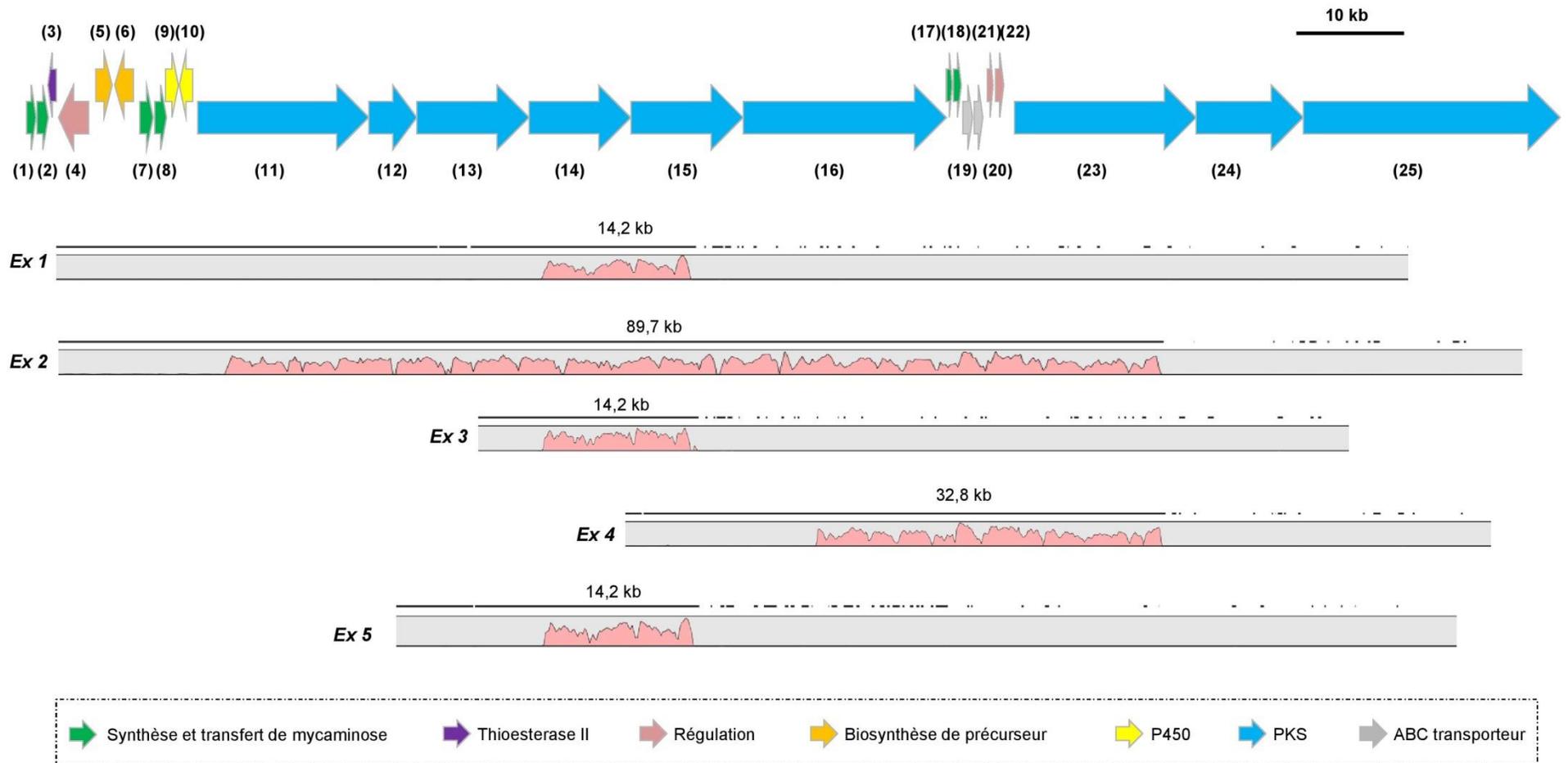


Figure R15 : Représentation des différentes AUD observées au sein du cluster de biosynthèse de la stambomycine. Les flèches de couleur indiquent les différents gènes du cluster (numérotés de 1 à 25). La fonction correspondant à chaque couleur est indiquée dans l'encadré au bas de la figure (d'après Laureti *et al.* 2010). Pour chaque exemple d'AUD sont représentées : (i) la taille en (kb), (ii) la position par rapport aux gènes du cluster et (iii) l'augmentation du taux de couverture des reads (les courbes roses dans les encadrés gris représentent la fluctuation du taux de couverture ; la valeur à gauche de l'encadré indique le taux de couverture maximal observé). Au-dessus de chaque encadré, la droite indique la présence ou l'absence de reads pour les loci correspondant.

Tableau R1 : Caractéristiques des AUD

Ex.	Taille	Position	Région homologue	
			Taille (pb)	Homologie (%)
1	14,2	7 719 700 - 7 733 900	1167	95
2	89,7	7 687 100 - 7 776 800	723	91
3	14,2	7 719 700 - 7 733 900	1167	95
4	32,8	7 744 000 - 7 776 800	740	96
5	14,2	7 719 700 - 7 733 900	1167	95

Ainsi, l'analyse des génomes démontre que la réparation de DSB au locus IKI-L engendre des réarrangements complexes et ce quel que soit le contexte génétique testé. L'identification des points de recombinaison suggère que la RH est un mécanisme de réparation de DSB prépondérant. Toutefois, des évènements de RI ont été décelés dans les trois contextes génétiques, suggérant ainsi l'existence d'un mécanisme *d'Alternative End-Joining* (AE-J). Un bilan sur la structure des différents génomes et du déroulement présumé des évènements de réparation est présenté en **figure R16**.

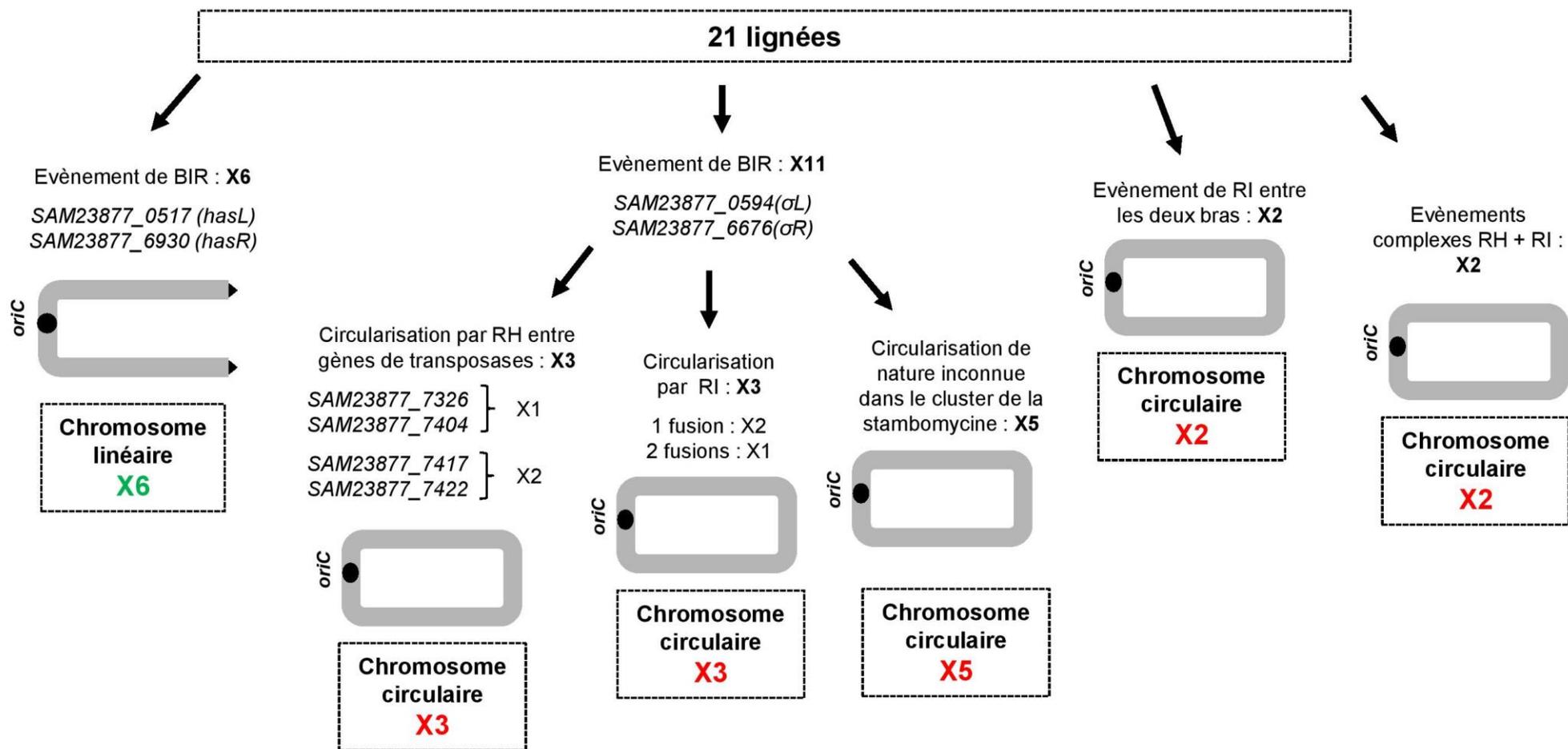


Figure R16 : Bilan de l'analyse des génomes séquencés

CHAPITRE 3 :

=

DISCUSSION GENERALE

I. Etude des acteurs tardifs de la recombinaison homologe chez *Streptomyces ambofaciens*

1. Rôle de RuvABC et RecG dans la migration et la résolution des jonctions de Holliday

Dans le cadre de notre projet visant à étudier l'impact des mécanismes de réparation des DSB sur la plasticité génomique, nous nous sommes focalisés sur l'identification des acteurs tardifs de la recombinaison homologe (RH) afin d'obtenir un mutant déficient pour ce mécanisme. Ainsi, nous nous sommes intéressés au complexe RuvABC et à l'hélicase RecG, connus pour être deux alternatives de l'étape post-synaptique de la RH chez *E. coli* (Lloyd et Sharples, 1993 ; Whitby *et al.*, 1993). Tandis que RuvAB et RecG catalysent la migration des jonctions de Holliday, l'endonucléase RuvC permet la résolution de la structure à quatre brins et le retour à une configuration initiale (Müller et West, 1994). Chez *S. ambofaciens*, en se basant sur deux approches axées sur des mesures de taux de recombinaison, nous avons démontré qu'un mutant $\Delta ruvABC \Delta recG$ est affecté pour la recombinaison conjugative, c'est-à-dire l'intégration d'un fragment d'ADN ou le remplacement d'un locus après transfert d'ADN, mais pas pour la recombinaison intra-chromosomique qui vise à déléter un locus grâce à un événement de RH entre deux séquences homologues. Toutefois, l'impact qu'entraînent ces mutations sur *Streptomyces* est faible en comparaison à d'autres bactéries telles qu'*E. coli* ou *B. subtilis*, ce qui nous a amené à suggérer que des protéines alternatives peuvent compenser les déficiences de RuvABC et RecG.

Une interrogation réside cependant dans la différence qu'il existe entre les résultats obtenus pour la recombinaison conjugative et la recombinaison intra-chromosomique. Lors de la conjugaison inter-générique avec *E. coli*, l'exogénote est transféré d'*E. coli* vers *Streptomyces* sous forme d'ADN simple-brin. Chez *E. coli*, il est connu que la persistance de l'ADN simple-brin au sein de la cellule entraîne l'induction du système SOS, un mécanisme de survie des bactéries en réponse aux lésions de l'ADN (**figure D1**).

En absence d'ADN simple-brin (d'origine exogène ou résultant de dommages à l'ADN), l'homodimère LexA, exprimé en quantité basale dans la cellule agit comme un répresseur sur une quarantaine de gènes ou opérons intervenant dans différents mécanismes de réparation d'ADN. Parmi ces loci, on retrouve notamment le gène *lexA* lui-même, ainsi que *recA* et *ruvAB*. En présence d'ADN simple-brin, la protéine RecA se fixe sur le fragment et la formation du complexe ADN simple-brin/RecA va activer cette dernière. Lorsque RecA est activée, son activité co-protéasique entraîne un clivage autocatalytique de

LexA qui l'empêche de former un dimère et entraîne une dérégulation des gènes SOS et par conséquent une augmentation de leur expression. Ainsi, en présence d'ADN exogène, la recombinaison interspécifique est favorisée grâce à l'augmentation de l'expression de *recA* et *ruvAB* (Matic *et al.*, 1995).

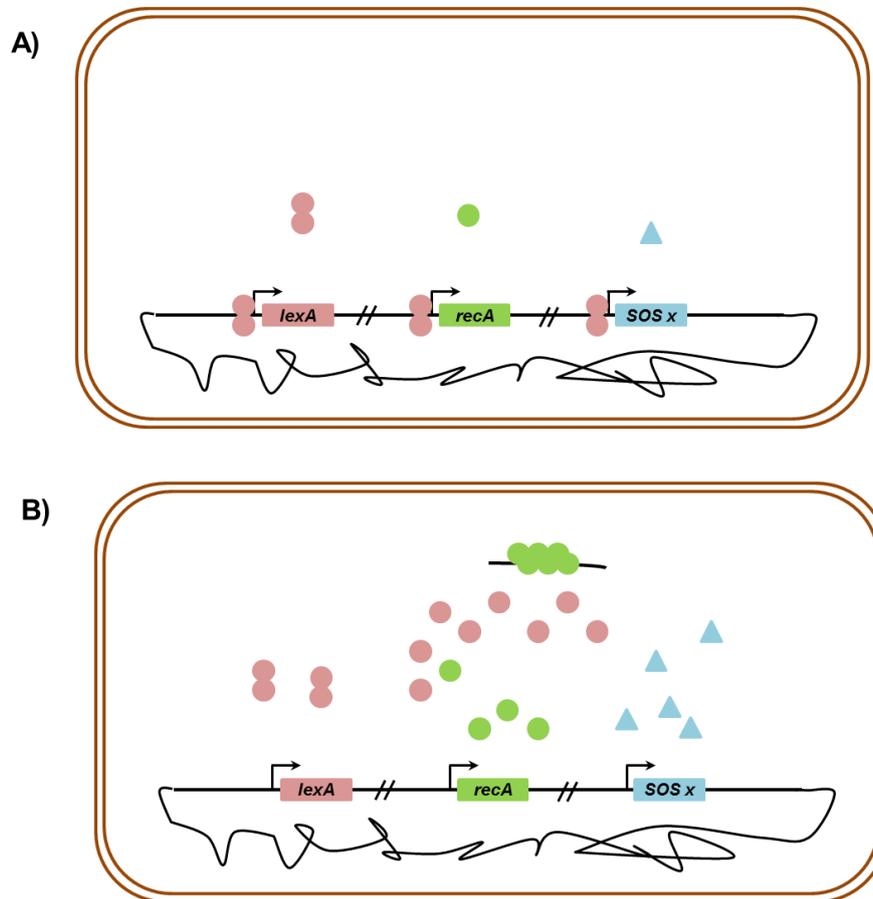


Figure D1 : Activation du système SOS d'*E. coli* par la présence d'ADN simple-brin. **A)** Dans une cellule en conditions optimales de croissance, les gènes SOS (représentés ici par *lexA* lui-même, *recA* et *SOSx* qui code un gène SOS quelconque) sont réprimés par la protéine LexA (dimère rose). Il existe toutefois une expression basale des gènes en présence du répresseur. **B)** En présence d'un fragment d'ADN simple-brin, la protéine RecA vient s'y fixer ce qui va entraîner son activation. Elle va alors interagir avec LexA et son activité co-protéasique entraîne un clivage autocatalytique de LexA et une dé-régulation des gènes SOS, et par conséquent une augmentation de l'expression des gènes de la RH et une facilitation de l'acquisition de matériel génétique exogène. Après intégration ou élimination du fragment, la protéine LexA va à nouveau prendre le dessus et réprimer les gènes SOS.

Si l'on extrapole ces résultats pour *Streptomyces*, il est envisageable que le transfert du BAC sous forme simple-brin de *E. coli* vers *Streptomyces* induise également l'expression

de *recA* par le biais d'un mécanisme SOS. Ainsi, l'augmentation de l'expression de *recA* stimulerait les premières étapes de la RH et renforcerait le besoin des enzymes intervenant dans les étapes tardives. En revanche, la recombinaison intra-chromosomique n'induirait pas le système SOS et la déficience de RuvABC et RecG pourrait être compensée par l'expression basale de protéines alternatives. A ce jour, peu d'informations sont toutefois disponibles sur un éventuel système SOS chez les *Streptomyces* qui possèdent toutefois une protéine LexA. Chez *S. avermitilis*, des analyses *in vitro* ont démontré que LexA peut se fixer au niveau d'une séquence promotrice en amont de *recA* et ainsi réguler son expression (Vierling *et al.*, 2001). Une séquence consensus similaire (**figure D2** ; source : www.regprecise.lbl.gov) a été retrouvée en amont d'une dizaine de gènes impliqués dans des mécanismes de réparation de l'ADN, dont *uvrA* qui intervient dans le mécanisme d'excision de nucléotides, ou encore *alkB* impliqué dans la réparation des dommages de l'alkylation. En revanche, cette séquence n'a été retrouvée ni en amont de *recG*, ni en amont de *ruvABC*.

L'activation des gènes de réparation de l'ADN pourrait toutefois se faire indépendamment de LexA chez les *Streptomyces*. Chez *M. tuberculosis*, il a été démontré qu'en plus du promoteur contrôlé par LexA, un second promoteur inductible par les dommages à l'ADN permet également la transcription de *recA* (Davies *et al.*, 2002). Par la suite il a également été démontré que même en absence du système de régulation LexA/RecA, la majorité des gènes impliqués dans des systèmes de réparation sont tout de même induits par les dommages de l'ADN (Rand *et al.*, 2003). En amont de ces différents gènes, une séquence consensus 5' TGTC-12nt-TAnnnT 3' a été identifiée (Gamulin *et al.*, 2004) et pourrait correspondre à une séquence promotrice. Chez *S. ambofaciens*, nous avons également pu identifier cette séquence en amont de plus d'une centaine de gènes, notamment *recA*, *ruvCAB* et *recG*, ainsi qu'un grand nombre de gènes hypothétiquement impliqués dans des mécanismes de réparation de l'ADN. Il est intéressant de noter que cette séquence consensus a également été retrouvée en amont de deux facteurs sigma qui pourraient ainsi être impliqués dans la transcription de ces différents gènes. Des expériences de mutagenèse ciblant ces deux loci pourraient être envisagées afin de déterminer leur importance au sein de la cellule.

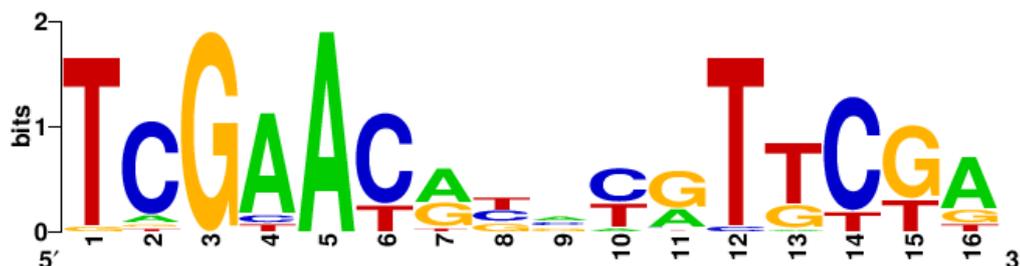


Figure D2 : Séquence consensus impliquée dans la régulation RecA/LexA dépendante chez les *Streptomyces*. La séquence logo a été réalisée à partir de 26 séquences trouvées en amont de gènes chez *S. avermitilis*, *S. coelicolor*, *S. griseus* et *S. scabiei*. Source : www.regprecise.lbl.gov

2. Possible redondance fonctionelle des protéines intervenant dans la migration et la résolution des jonctions de Holliday

Afin d'envisager une identification de protéines alternatives pour l'étape post-synaptique de la RH, le génome de *S. ambifaciens* a été analysé à la recherche de candidats potentiels. Cette analyse a révélé notamment le gène *SAM23877_3744* qui code un homologue de la protéine DnaB d'*E. coli*. En effet, des analyses *in vitro* effectuées sur cette hélicase hexamérique d'*E. coli* ont démontré qu'elle peut se lier à une extrémité 5' d'un bras de la JH et catalyser sa migration après encerclement des deux brins (Kaplan *et al.*, 2012). De la même manière, les gènes *SAM23877_4613* et *SAM23877_4970* codent respectivement UvrD1 et UvrD2 qui sont des homologues d'UvrD d'*E. coli*, une protéine impliquée dans les systèmes de réparation de mésappariements et de nucléotides aberrants par excision de nucléotides (NER). Chez *E. coli*, quand les étapes tardives de la RH sont bloquées (exemple d'une mutation pour *ruvABC*), la délétion d'*uvrD* est létale (Magner *et al.*, 2007). Cette particularité pourrait en partie être expliquée par la faculté d'UvrD à dissocier les nucléofilaments formés par RecA et ainsi supprimer les intermédiaires de la RH qui s'accumulent (Veauté *et al.*, 2005). Cependant, il a également été démontré *in vitro* qu'UvrD peut dérouler des JH et pourrait donc avoir un rôle redondant avec RuvAB et RecG (Carter *et al.*, 2012).

A un degré moindre, la migration des JH peut également être assurée par certaines protéines dépourvues d'activité hélicase. En plus de son activité de recherche d'homologie et d'échange de brins, l'ATPase RecA peut également intervenir dans la migration des JH, mais son mode d'action qui reste méconnu est relativement inefficace en comparaison des hélicases RuvAB et RecG (Adams et West, 1996 ; Persky et Lovett 2008). Un autre candidat pourrait être un paralogue de *recA* connu sous les noms de *sms* ou *radA* (*SAM23877_4210*)

chez *Streptomyces* et ubiquiste dans le monde bactérien (Lin *et al.*, 2006). Chez *E. coli*, la mesure de fréquences de recombinaison suggère un rôle redondant de RadA/Sms avec les autres acteurs de la migration des JH (Beam *et al.*, 2002). En effet, comme pour des souches $\Delta ruvAB$ et $\Delta recG$, un mutant $\Delta radA/sms$ est faiblement impacté pour la RH tandis qu'un triple mutant $\Delta radA/sms \Delta recG \Delta ruvABC$ est sévèrement déficient pour la RH, d'une manière comparable à un mutant $\Delta recA$ (Beam *et al.*, 2002). Récemment, une étude a confirmé que RadA/Sms peut catalyser la migration des JH mais étant donné qu'elle ne dispose pas d'activité hélicase, il est suggéré que son mode d'action soit différent de RuvAB et RecG (Cooper et Lovett, 2016).

Pour la résolution des JH, au moins deux candidats alternatifs à RuvC peuvent être proposés. Comme chez *E. coli*, un *locus* associé aux CRISPR de type I qui code le gène spécifique *cas3* ainsi que les gènes ubiquistes des systèmes CRISPR *cas1* et *cas2* ont été identifiés chez *S. ambofaciens*. Bien que le rôle principal des systèmes CRISPR ne soit pas lié à la RH, des analyses *in vitro* ont démontré que la nucléase Cas1 d'*E. coli* peut procéder à la résolution des JH (Babu *et al.*, 2011). Toutefois, contrairement à RuvC, Cas1 ne cible pas de séquence particulière ce qui suggère que son rôle dans la résolution des JH serait opportuniste plutôt que programmé. Enfin, le second candidat qui peut être suggéré paraît être l'alternative la plus probable à RuvC : Aravind et ses collaborateurs ont identifié la famille de protéines YqgF qui partage des homologies structurales avec les protéines de la famille de RuvC (Aravind *et al.*, 2000). Très récemment, des analyses *in vitro* effectuées sur RuvX, le membre de la famille YqgF chez *Mycobacterium*, ont révélé que cette protéine a une forte affinité pour les JH et peut catalyser leur clivage (Nautiyal *et al.* 2016). *Mycobacterium* et *Streptomyces* appartenant tous deux à la famille des actinobactéries, l'analyse du génome de *S. ambofaciens* a permis de détecter un homologue de RuvX (*SAM23877_1552*). Cette protéine étant un sérieux candidat pour la résolution des JH, il serait intéressant de muter le *locus* et tester son implication dans la RH.

II. Le NHEJ chez *Streptomyces*

Les travaux initiés il y a quelques années au sein du laboratoire (ANR Streptoflux 2007-2011) et poursuivis au cours de ma thèse ont permis de mettre en évidence l'existence d'un mécanisme de NHEJ intervenant dans la réparation des DSB chez les *Streptomyces*. Chez les eucaryotes, le NHEJ est ubiquiste et nécessite la présence de nombreux acteurs. Le constat est différent pour les procaryotes, où seuls 25% des organismes possèdent au moins les deux gènes suffisants pour un NHEJ fonctionnel : *ku* et *ligD* (McGovern *et al.*, 2016). Ainsi, l'identification d'un grand nombre de gènes potentiellement impliqués dans le NHEJ chez une bactérie telle que *Streptomyces* nous amène à nous poser la question du rôle d'un tel mécanisme et de ces acteurs dans la biologie de l'organisme.

1. Importance du NHEJ dans le cycle cellulaire de la bactérie

Le NHEJ n'est présent que chez une fraction d'espèces bactériennes, ce qui suggère qu'il n'a pas la même importance pour les procaryotes que pour les eucaryotes. En conditions de laboratoire, des bactéries telles que *B. subtilis* (Weller *et al.*, 2002) ou encore *M. tuberculosis* (Gong *et al.*, 2005) mutées pour le NHEJ ne présentent pas de défaut de croissance. De la même manière, les colonies de différentes souches délétées pour un ou plusieurs gènes NHEJ-like chez *S. ambofaciens* ne montrent pas de différences phénotypiques en comparaison avec une souche sauvage. Des analyses bio-informatiques effectuées sur les génomes bactériens séquencés ont démontré que les espèces bactériennes possédant des gènes NHEJ-like ne sont pas phylogénétiquement liées (**figure D3**), ce qui plaide en faveur d'une acquisition par transfert horizontal ou de pertes du mécanisme ancestral au sein de certains phyla (Rocha *et al.*, 2005 ; Bowater et Doherty, 2006). L'ensemble des bactéries pour lesquelles un mécanisme de NHEJ a été mis en évidence partagent une caractéristique commune : elles passent une grande partie de leur cycle cellulaire en phase stationnaire ou de dormance (Pour revue : Bowater et Doherty, 2006 ; Pitcher *et al.*, 2007a ; Brissett et Doherty, 2009). C'est notamment vrai pour le pathogène *M. tuberculosis*, capable de survivre plusieurs années en dormance à l'intérieur d'un macrophage, mais également pour *B. subtilis* qui forme une endospore lorsque les nutriments deviennent limitants. Dans les deux cas, les cellules sont exposées à de nombreux stress et subissent l'attaque de facteurs physico-chimiques pouvant engendrer des lésions sur l'ADN incluant des DSB. Du fait de l'état quiescent des cellules et donc de l'absence de réplication, il est suggéré que l'intervention de la RH pour la réparation d'une

DSB soit fortement limitée, et que la présence d'un mécanisme de NHEJ ait été sélectionnée pour pallier cela. Cette hypothèse est étayée par la démonstration que des souches de *Mycobacterium* délétées pour le NHEJ sont plus sensibles que la souche sauvage à l'exposition aux rayons ionisants lorsqu'elles sont en phase stationnaire mais pas en phase exponentielle de croissance (Pitcher *et al.*, 2007b). De la même manière, des spores de *B. subtilis* déficientes pour le NHEJ ont une sensibilité accrue à l'exposition aux rayons ionisants ou à une sécheresse extrême (Moeller *et al.*, 2007). Le lien entre le stade « spore » chez *B. subtilis* et la réparation NHEJ est renforcé par le contrôle de l'opéron *ykoVU* (*ku* et *ligD*) par σ^G , un facteur sigma qui contrôle également l'expression des gènes intervenant dans la formation de l'endospore (Wang *et al.*, 2006).

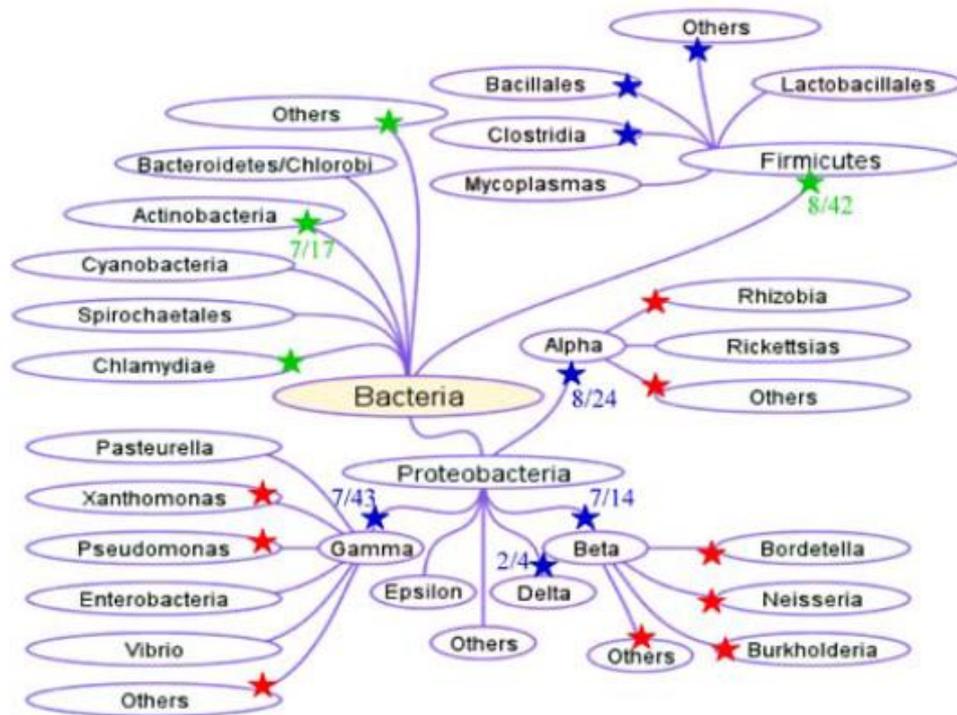


Figure D3 : Distribution phylogénique des protéines Ku chez les bactéries. Les étoiles indiquent les phyla (vert), classes/ordres (bleu) et familles (rouge) qui contiennent au moins un homologue de *ku* (Bowater et Doherty, 2006)

Comme cela a été décrit dans l'introduction, le cycle de développement de *Streptomyces* alterne croissance mycélienne et phase de sporulation. Durant la phase de sporulation, il est admis qu'une seule copie du chromosome est présente au sein de chaque cellule. Ainsi, la réparation par RH serait limitée et le NHEJ pourrait jouer un rôle prépondérant dans la réparation des DSB. L'exposition de spores mutées pour les gènes NHEJ-like à différentes doses de rayons ionisants nous a permis de constater que ces

dernières sont affectées dans la réponse aux dommages de l'ADN, confortant ainsi cette hypothèse.

2. Caractérisation des gènes NHEJ-like chez les *Streptomyces*

2.1 Implication dans le NHEJ

Avantages d'un mécanisme NHEJ à trois Ku

Une des grandes différences entre le NHEJ de *Streptomyces* et celui des autres bactéries réside dans l'organisation des gènes NHEJ-like et de leur forte variabilité interspécifique. Ainsi, *S. ambofaciens* dispose de trois gènes qui codent des protéines homologues de *ku*. Le gène *kuA* est co-localisé avec le gène *polK* dans la région « core » du chromosome et est conservé dans l'ensemble des génomes de *Streptomyces*. En revanche *kuB* et *kuC* sont sporadiques au sein des espèces. Le maintien de ces gènes au sein du génome de *Streptomyces* implique qu'ils contribuent à la valeur sélective de la bactérie. Dans le cas contraire, ces gènes auraient été perdus par accumulation de mutations ou par des événements de recombinaison. Les expériences d'exposition aux rayons électroniques nous ont révélé que les trois protéines Ku sont impliquées dans la réponse aux dommages à l'ADN. Toutefois, KuA semble être la protéine Ku centrale du NHEJ chez les *Streptomyces* pour au moins 3 raisons. En effet, (i) le gène est conservé dans l'ensemble des espèces (à l'exception de *S. vietnamensis*), (ii) un mutant ΔkuA est sensiblement plus affecté que les mutants ΔkuB et ΔkuC dans la réponse aux rayons électroniques, et (iii) l'absence de *kuA* entraîne une diminution d'un facteur 100 de la réparation par NHEJ dans la région centrale du chromosome.

Dans les modèles bactériens classiques, *Mycobacterium*, *Bacillus subtilis*, *Pseudomonas aeruginosa*, un seul gène codant une protéine Ku a été identifié. Cependant, comme pour *Streptomyces*, certaines bactéries possèdent plusieurs exemplaires de *ku*. C'est notamment le cas des α -protéobactéries *Agrobacterium tumefaciens* et *Sinorhizobium meliloti* qui disposent respectivement de 3 et 4 gènes *ku*. De manière similaire aux résultats obtenus pour *Streptomyces*, les différents mutants *ku* de *Sinorhizobium* présentent une sensibilité accrue aux rayons ionisants en comparaison d'une souche sauvage (ii) et les analyses transcriptionnelles suggèrent que l'expression des gènes est induite dans des conditions spécifiques (Kobayashi *et al.*, 2008).

L'hypothèse principale sur la multiplicité des *ku* chez les *Streptomyces* est la suivante : le gène *kuA* ancestral conservé au sein de tous les génomes interviendrait principalement dans la réparation des DSB dans les spores tandis que la présence des gènes *kuB* et *kuC* résulterait d'une acquisition par THG. Ces derniers seraient exprimés dans des conditions particulières dans lesquelles ils contribueraient à améliorer l'adaptabilité de la bactérie. L'analyse phylogénétique de la distribution des *ku* ne nous permet cependant pas d'affirmer que la présence de *kuB* et *kuC* résulte d'une acquisition par THG. Leur absence dans certains génomes pourrait résulter d'une perte du locus ayant notamment pu survenir suite à un réarrangement ou à une modification de l'écosystème entraînant une obsolescence de la fonction du gène. Toutefois, il est à noter que *kuB* et *kuC* sont localisés sur des plasmides dans certains génomes ce qui plaide en faveur des THG. Les analyses d'expression de ces trois gènes tout au long du cycle cellulaire de la bactérie montrent une accumulation des transcrits en phase de sporulation pour *kuA* mais pas pour *kuB* et *kuC* ce qui suggère une régulation différente.

La multiplicité des *ku* pourrait également entraîner une augmentation de la compétitivité du NHEJ par rapport aux autres mécanismes de réparation des DSB (RH). En effet, dans les conditions adéquates, l'expression des trois gènes *ku* augmenterait la quantité de protéines Ku au sein de la cellule et ainsi la probabilité de prise en charge des DSB par le NHEJ. Comme nous avons pu le constater au travers des expériences de génération de DSB ciblées, le NHEJ est un mécanisme mutagène pouvant entraîner la délétion ou l'intégration d'ADN au niveau du site de cassure. Ainsi une prise en charge accrue des DSB par le NHEJ favoriserait l'intégration à court terme de matériel génétique exogène et augmenterait la capacité d'adaptation de la bactérie à son environnement. Si une corrélation entre l'augmentation de l'adaptabilité et la multiplicité des *ku* n'a jamais été démontrée, des informations peuvent en revanche être tirées d'organismes ne disposant pas de NHEJ. Ainsi, la levure *Lachancea kluyveri*, seul organisme eucaryote chez lequel l'absence de NHEJ a à ce jour été mise en évidence, présente un nombre de réarrangements génomiques significativement plus bas que les autres espèces de la famille des Saccharomycètes. Les auteurs suggèrent que la perte du NHEJ est responsable du faible taux de recombinaison au sein de cette levure (Gordon *et al.* 2011).

LigC et LigD : deux ligases pour la réparation des DSB

La ligation est une étape primordiale lors de la réparation des DSB par le mécanisme de NHEJ. En effet après traitement des extrémités, l'intervention d'une ligase permet de

suturer les deux extrémités de la cassure entre elles et ainsi d'achever le processus de réparation. Chez *Mycobacterium*, deux protéines interviennent pour cette étape : LigD et à un degré moindre LigC. En effet, il a été démontré chez *Mycobacterium smegmatis* que lors de la réparation de DSB à bouts francs, l'efficacité de réparation diminuait de 65 fois en absence de LigD (Gong *et al.*, 2005). Lorsque la délétion de *ligD* est accompagnée de celle de *ligC*, la diminution de l'efficacité est aggravée de 24 fois par rapport au simple mutant $\Delta ligD$. En revanche, la délétion de *ligC* seule n'a aucun impact sur l'efficacité du NHEJ ce qui suggère que la protéine a un effet *back-up* de LigD (Gong *et al.*, 2005). Chez *Streptomyces*, si l'on se fie à la distribution des gènes *lig* dans les différents génomes, *ligC* pourrait être le gène codant la ligase principale du NHEJ puisqu'il (i) est conservé dans l'ensemble des génomes et (ii) est co-localisé avec *polR*, un gène homologue du domaine polymérase de LigD (Pol-Dom), tandis que le domaine ligase de *ligD* est uniquement présent dans 29 des 38 génomes. La sensibilité des gènes de chacun des deux gènes aux rayons électroniques confirme que les ligases ont un rôle dans la réponse aux dommages de l'ADN. La génération de DSB ciblées dans la région centrale du chromosome montre une absence de réparation par NHEJ dans un contexte $\Delta ligCD$, ce qui démontre qu'au moins l'une de ces deux ligases joue un rôle majeur dans ce mécanisme. Les deux ligases pourraient contribuer ensemble au NHEJ avec une spécificité pour des extrémités de nature différente. Afin d'étayer cette hypothèse, il serait nécessaire de compléter nos expériences en testant la conséquence de génération de DSB à extrémités franches, 3' sortantes et 5' sortantes dans des souches mutées pour *ligC* ou *ligD*.

Rôle de polK et polR dans le NHEJ

Les gènes *polK* et *polR* sont deux homologues du domaine polymérase de *ligD* (Pol-Dom). Les analyses *in silico* ont révélé qu'ils sont conservés dans l'ensemble des génomes (à l'exception de *S. vietnamensis* pour *polK*) et bien qu'ils ne soient pas en opéron, ils sont toujours co-localisés avec respectivement *kuA* pour *polK* et *ligC* pour *polR*. La sensibilité des mutants $\Delta polR$ et $\Delta polK$ aux rayons électroniques montre que ces gènes ont un rôle dans la réponse aux dommages de l'ADN et pourraient ainsi être impliqués dans le NHEJ. Toutefois, la conservation des deux loci suggère une absence de redondance fonctionnelle. Dans le NHEJ, le rôle enzymatique d'une polymérase se caractérise soit par : (i) *fill-in*, c'est-à-dire par comblement de lacunes simple-brin formées par jonction d'extrémités rendues compatibles par la présence de microhomologies ou (ii) synthèse aléatoire de nucléotides aux extrémités de la cassure. Chez *M. smegmatis*, la mutation de Pol-Dom entraîne l'abolition

de l'ajout de nucléotides aux extrémités de la cassure, mais n'impacte pas le *fill-in*, ce qui suggère l'action de polymérases distinctes (Aniukwu *et al.*, 2008). Il pourrait en être de même chez *Streptomyces* pour PolK et PolR.

Chez *Mycobacterium* comme chez *Streptomyces*, l'intervention de la polymérase n'est pas indispensable pour la réparation d'une DSB par NHEJ et est plus ou moins importante en fonction de la nature des extrémités (Aniukwu *et al.* 2008). Chez *Streptomyces*, l'analyse de la réparation d'extrémités 3' sortantes engendrées par le système *I-SceI* dans un contexte sauvage a révélé une intervention d'une activité polymérase dans 40% des cas, avec majoritairement du *fill-in* (18 cas sur 26) et plus rarement l'ajout d'un nucléotide unique aux extrémités (8 cas sur 26). Ces résultats soulignent l'impact faible de l'activité polymérase relativement à la sensibilité conférée par la mutation double *polRK* dans la réponse aux dommages de l'ADN. Cette différence pourrait révéler que l'activité polymérase n'est pas la fonction principale des protéines Pol chez *S. ambofaciens*. En effet, chez *Mycobacterium*, en plus de son activité polymérase, Pol-Dom permet le recrutement de LigD par Ku (Pitcher *et al.*, 2005). Ainsi, la délétion des gènes *pol* chez *S. ambofaciens* pourrait entraîner une déficience du NHEJ non pas par l'absence d'activité polymérase, mais par une incapacité à recruter les ligases. En revanche, la forte conservation de deux protéines Pol indépendantes serait plutôt liée à la présence de deux activités polymérase distinctes entraînant un impact différent sur la plasticité génomique. Des expériences de double hybride en cours au laboratoire pourraient permettre de tester ces hypothèses.

Une activité nucléasique non identifiée

S'il est admis que la réparation de DSB par NHEJ chez les bactéries s'accompagne généralement de délétions pouvant atteindre plusieurs centaines de bases de part et d'autre du site de cassure, la ou les protéines responsable(s) de cette activité restent toutefois à identifier. Le domaine nucléase de LigD (Nuc-Dom) semble à première vue être le candidat le plus probable pour remplir cette fonction. Cependant, il est absent chez la protéine LigD de *B. subtilis* (constituée uniquement d'un domaine polymérase et d'un domaine ligase), bactérie chez laquelle des délétions liées au NHEJ sont également observées (de Vega, 2013). Chez *M. tuberculosis*, il a été démontré *in vitro* que la protéine LigD possède une activité exonucléasique de 3' vers 5' qui serait dépendante du domaine nucléase, puisque lorsque ce dernier est délété, cette activité n'est plus détectable (Della *et al.*, 2004). Pour autant, aucune activité *in vitro* n'a pu être caractérisée lorsque le domaine Nuc-Dom est testé seul (Pitcher *et al.*, 2005). De plus, l'induction de cassures ciblées de différentes natures dans une souche

mutée pour Nuc-Dom chez *Mycobacterium* ne montre aucune différence par rapport à une souche sauvage dans l'efficacité de réparation ainsi que dans la taille des délétions entraînées par le NHEJ (Aniukwu *et al.*, 2008). Malgré plusieurs études sur Nuc-Dom, son rôle dans le NHEJ reste peu connu. Chez *S. ambofaciens*, les résultats montrent des délétions de nucléotides de part et d'autre du site de cassure lors des événements de réparation par NHEJ alors qu'aucun gène homologue de Nuc-Dom n'a été identifié, ce qui suggère l'intervention d'une ou plusieurs nucléases encore inconnues.

Lors de l'analyse des cicatrices de réparation laissées par le NHEJ, deux exemples se sont avérés particulièrement intéressants puisqu'ils révèlent qu'une petite partie du *locus* délété lors du traitement des extrémités a été réinséré lors de la réparation. L'intégration de ces fragments de respectivement 42 et 83 nucléotides suggère que l'enzyme n'agit pas de façon processive mais qu'elle dispose d'une activité endonucléolytique qui produit des petits fragments. Ce mode d'action a notamment été décrit pour le complexe hélicase-nucléase RecBCD qui intervient dans l'étape pré-synaptique de la RH chez *E. coli*. La progression du complexe le long de la molécule qui se fait grâce à l'activité hélicase de RecB et RecD est suivie du clivage endonucléolytique de l'ADN par l'intermédiaire du domaine nucléase de RecB (pour revue : Dillingham et Kowalczykowski, 2008). Toutefois, l'idée d'une intervention de RecBCD dans le NHEJ des actinomycètes ne paraît pas envisageable. En effet, bien que présent chez *M. tuberculosis*, la mutation de ce *locus* n'affecte ni l'efficacité ni la nature des délétions liées au NHEJ (Aniukwu *et al.*, 2008). Chez les *Streptomyces*, le *locus* est absent de l'ensemble des génomes disponibles. Un second complexe hélicase-nucléase nommé AdnAB a également été identifié chez *Mycobacterium* (Sinha *et al.*, 2009). Des analyses *in vitro* ont permis d'établir un modèle présentant l'intervention d'AdnAB dans l'étape pré-synaptique de la RH (Unciuleac *et al.*, 2009) (figure D4). Bien que son mode de fonctionnement soit compatible avec l'hypothèse formulée précédemment pour la nucléase, l'implication d'AdnAB dans le NHEJ chez *Mycobacterium* semble peu probable puisqu'il (i) a été démontré *in vitro* qu'il existe une compétition entre Ku et AdnAB pour la reconnaissance des DSB, et (ii) que la fixation de Ku empêcherait ainsi la résection des extrémités par AdnAB (Sinha *et al.*, 2009). Au laboratoire, un homologue d'*adnAB* a été identifié chez *S. ambofaciens* (Zhang *et al.* 2014). Contrairement à *Mycobacterium*, le *locus* est essentiel pour la survie de la bactérie. Etant donné qu'il est possible d'obtenir un mutant Δ *recA* déficient pour la RH chez *S. ambofaciens*, l'essentialité d'*adnAB* ne réside pas dans son implication dans la RH, ce qui suggère que le complexe a un rôle pléiotrope. Ainsi, au vu des données

dont nous disposons actuellement sur *adnAB* de *S. ambofaciens*, il n'est pas possible d'exclure son implication à la fois dans la RH et dans le NHEJ.

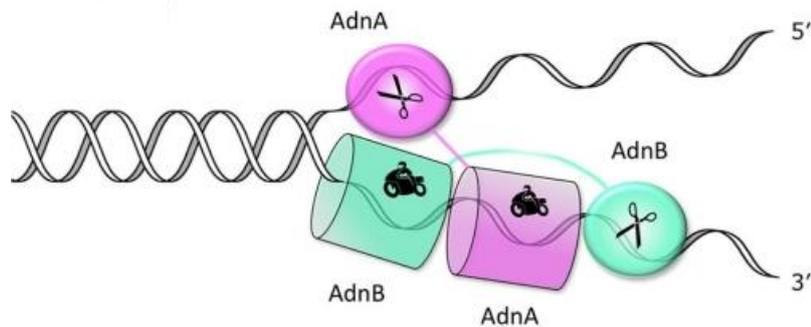


Figure D4 : Modèle du mode d'action d'AdnAB pour la résection des extrémités. AdnB (turquoise) déroule le double-brin d'ADN ce qui permet à AdnA (violet) de suivre sa piste en tandem. L'activité nucléase d'AdnA permet la dégradation de 5' vers 3' tandis que l'activité nucléase d'AdnB permet la dégradation de 3' vers 5' (Unciuleac *et al.*, 2009)

Chez *Bacillus*, il a été suggéré que le rôle de nucléase pourrait incomber à SbcD (Ayora *et al.*, 2011 ; de Vega, 2013), une protéine qui possède une activité exonucléasique 3' vers 5' ainsi qu'une activité endonucléasique sur les deux brins de l'ADN. (Connely *et al.*, 1997). SbcD forme un complexe avec la protéine SbcC connue pour intervenir dans la réparation des DSB (Mascarenhas *et al.*, 2006). Elle appartient également à la même famille que la protéine eucaryote Mrx11 impliquée dans le NHEJ chez la levure et qui interagit avec Ku80 (Daley *et al.*, 2005). Chez les *Streptomyces*, nous avons pu identifier des homologues pour SbcC et SbcD (chez *S. ambofaciens* : SAMR23877_1380 et SAMR23877_1381) codées par deux gènes qui forment probablement un opéron. Ce complexe nucléasique pourrait être un bon candidat pour le NHEJ chez les *Streptomyces* et il faudrait d'une part le muter pour voir son impact dans la réparation des DSB et d'autre part tester son interaction avec les autres protéines du NHEJ.

2.2 Implication des gènes NHEJ-like dans d'autres mécanismes de réparation de l'ADN

L'existence de gènes NHEJ-like multiples et leur implication respective dans la réponse aux dommages de l'ADN pourrait également résider dans l'intervention de certains de ces gènes dans des mécanismes de réparation de l'ADN autre que le NHEJ.

La mutagénèse en phase stationnaire

Récemment, les gènes du NHEJ chez *Pseudomonas putida* ont été impliqués dans la mutagénèse en phase stationnaire ou mutagénèse adaptative (Paris *et al.*, 2015). Durant la phase stationnaire, la mutagénèse adaptative engendre au sein de la population des mutants capables de surmonter les conditions nutritives limitantes et d'initier un nouveau cycle végétatif. Si le rôle des acteurs du NHEJ dans ce phénomène reste à préciser, la délétion de *ku* et *ligD* entraîne une déficience pour la mutagénèse adaptative. Bien que ce phénomène n'ait pas été décrit chez les *Streptomyces*, l'intervention des gènes NHEJ-like dans des mécanismes de mutagenèse pourrait favoriser l'adaptation en réponse aux carences nutritives rencontrées dans les écosystèmes occupés.

Le mécanisme d'excision de bases

Au sein de la cellule, la formation de sites abasiques (absence de bases sur certains nucléotides, également appelés sites AP) est un dommage fréquent, particulièrement récurrent lorsque les bactéries sont exposées à des conditions environnementales restrictives. Dans ces cas-là, l'intervention du mécanisme d'excision de base (BER) permet de réparer les sites AP. Dans ce processus, une enzyme appelée AP lyase ou AP endonucléase permet le clivage du désoxyribose. Puis l'intervention d'une polymérase permet de remplir l'espace libéré en synthétisant une nouvelle base en utilisant comme matrice le brin complémentaire. Enfin une ligase suture le brin réparé. Chez *B. subtilis*, il a récemment été démontré que les protéines Ku et LigD possèdent une activité AP-lyase et pourraient ainsi être impliquées dans le mécanisme de BER (de Ory *et al.*, 2014 ; de Ory *et al.*, 2016). Les activités polymérase et ligase de LigD finaliseraient le processus. Cette activité a également été démontrée chez *P. aeruginosa*, bactérie phylogénétiquement éloignée de *B. subtilis*, ce qui suggère qu'elle puisse être conservée chez l'ensemble des bactéries possédant un mécanisme NHEJ. Ainsi, il pourrait en être de même chez les *Streptomyces*.

Les Ku et le maintien des télomères

Contrairement à la majorité des bactéries, le chromosome de *Streptomyces* présente une conformation linéaire, avec des séquences terminales inversées répétées. Les queues 3' des télomères forment une structure secondaire en forme de Y aux extrémités desquelles viennent se fixer les protéines TAP (*terminal associated protein*) et TP (*terminal protein*) qui

participent à la réplication des télomères (Huang *et al.*, 1998). Les télomères peuvent être la cible d'activité nucléasique et de fusion entre extrémités chromosomiques et doivent donc être protégées. Chez les eucaryotes, le rôle paradoxal de Ku dans le maintien et la protection des télomères est prouvé de longue date (Pour revue : Fell et Schild-Poulter, 2015), et semble être indépendant du NHEJ puisque la délétion de la ligase NHEJ n'impacte pas l'intégrité des télomères. Toutefois, la manière dont Ku se fixe *in vivo* aux télomères reste inconnue. Bien que des expériences *in vitro* aient démontré la faculté de Ku à se fixer directement aux télomères, le recrutement de Ku pourrait également se faire par interaction avec des protéines associées aux télomères. Chez *Streptomyces*, un rôle similaire des Ku dans la protection des télomères pourrait également être envisageable. Ainsi, une hypothèse possible serait une interaction de Ku avec les protéines terminales, changeant ainsi la conformation de Ku ce qui permettrait une distinction entre son rôle dans le NHEJ et celui dans la protection des télomères. Des expériences d'interaction protéine-protéine permettraient de tester cette hypothèse.

III. Prise en charge des cassures double-brin le long du chromosome

Un des objectifs majeur de cette thèse a été d'introduire des DSB à des loci ciblés du chromosome de *S. ambofaciens* afin de pouvoir étudier leur prise en charge par les mécanismes de réparation et d'analyser leur impact sur la plasticité génomique. Cette expérience a été réalisée dans trois contextes génétiques distincts : un contexte sauvage (WT), un contexte déficient pour *kuA*, seul gène *ku* conservé, et un dernier contexte délété de *ligC* et *ligD* donc dépourvu des homologues des deux ligases qui interviennent dans le NHEJ chez *Mycobacterium*.

1. Conséquences de la réparation de DSB sur le chromosome

1.1 Des modifications mineures liées au NHEJ dans la région core

Le traitement des extrémités se caractérise par des délétions et/ou synthèses de nucléotides

Dans la région centrale du génome, très peu d'évènements de réparation mutagène ont été observés. En effet, moins d'1% de clones (82 sur 10 000 testés) présentent dans un contexte sauvage des remaniements entraînant une perte de la résistance à la kanamycine. Aucun de ces événements n'a été observé dans un contexte $\Delta ligCD$, ce qui démontre l'implication du NHEJ. Dans un contexte ΔkuA , l'obtention d'un évènement de réparation par recombinaison illégitime (RI) suggère qu'au minimum un acteur qui pourrait être KuB ou KuC (ou les 2) peut suppléer KuA, mais avec une fréquence moindre. L'analyse de cicatrices laissées après réparation a permis d'établir les caractéristiques du mécanisme de NHEJ chez *Streptomyces*.

L'ensemble des évènements de réparation analysés présentent des caractéristiques typiques du mécanisme de NHEJ, à savoir des délétions de part et d'autre des extrémités de la cassure et plus rarement une synthèse de quelques nucléotides pouvant probablement faciliter la ligation. Dans les exemples analysés, l'ampleur des délétions peut varier de quelques nucléotides à plusieurs dizaines de kb. Ainsi, la taille des délétions engendrées par le NHEJ chez *S. ambofaciens* pourrait être plus importante que chez *Mycobacterium*, où elle est généralement inférieure à 1 kb. (Stephanou *et al.*, 2007 ; Aniukwu *et al.*, 2008 ; Gupta *et al.*, 2011). Dans une étude récente, il a été suggéré que la taille importante des délétions pourrait être liée à l'absence d'une protéine multi-fonctionnelle LigD (Tong *et al.*, 2015). En effet, après complémentation hétérologue avec un gène *ligD* (codant les trois domaines) issu d'une bactérie du genre *Nocardia*, la taille des délétions observées après réparation des DSB

a diminué. Ainsi, il est envisageable que la présence des trois activités au sein d'une même protéine augmente la fidélité du mécanisme de NHEJ.

Le NHEJ permet l'insertion d'ADN au site de cassure

Toutefois, la caractéristique la plus remarquable qu'il nous a été donné d'observer pour 3 des clones analysés, concerne l'insertion de fragments d'ADN au site de réparation de la cassure. Dans deux des trois cas, la séquence insérée provient de la région chromosomique qui est délétée *in fine*. La présence de ce fragment pourrait donc résulter de l'action endonucléolytique de la nucléase. Dans le troisième cas, la séquence de 30 nucléotides insérée est identique à 100% avec une région contenue dans les répétitions terminales inversées du chromosome. Bien que dans notre dispositif expérimental, la séquence insérée au point de réparation ne puisse dériver que du chromosome de la souche étudiée, ces données suggèrent que le mécanisme de NHEJ permette l'intégration d'ADN exogène par transfert horizontal. Chez la levure *Schizosaccharomyces pombe* il a été montré que le NHEJ favorisait l'intégration d'ADN mitochondrial au sein du chromosome (Decottignies, 2005). En revanche, chez les procaryotes, le lien entre acquisition de matériel génétique extra-chromosomique et NHEJ reste peu connu, et des expériences visant à confirmer cette hypothèse sont envisagées au sein de l'équipe (voir perspectives). Nous avons vu dans l'introduction que de nombreux évènements de THG avaient pu être mis en évidence chez les *Streptomyces*. Contrairement à la RH, le NHEJ permettrait d'intégrer du matériel génétique même en absence de séquences homologues, ce qui en ferait un mécanisme évolutif très important. Ainsi, le NHEJ pourrait être au cœur d'une dynamique de flux de gènes au sein du génome en favorisant à la fois la délétion et l'insertion de matériel génétique.

1.2 Des réarrangements majeurs dans les bras du chromosome

Lors des analyses de génome, nous avons constaté que des phénomènes de remplacement de bras, d'amplification d'ADN, et de circularisation du chromosome joueraient un rôle primordial pour la réparation de DSB au site IKI-L. Les paragraphes suivants permettront de discuter des mécanismes impliqués et de leur impact sur la structure du génome.

1.2.1 La recombinaison homologe joue un rôle prépondérant dans les réarrangements génomiques observés

Le remplacement de bras, un évènement prédominant pour l'initiation de la réparation

En amont du site IKI-L, deux loci qui seraient particulièrement impliqués dans la réparation des DSB RH-dépendantes ont été mis en évidence. En effet, *SAM23877_0517* (*hasL*) et *SAM23877_0594* (σL) présentent la particularité d'avoir une copie identique à plus de 98% disposée en orientation inverse sur le bras droit, à savoir respectivement *SAM23877_6930* (*hasR*) et *SAM23877_6676* (σR). Dans 17 des 21 génomes analysés par séquençage, l'un ou l'autre de ces deux couples de gènes serait responsable de l'initiation de la réparation de la cassure grâce à un évènement de remplacement du bras gauche endommagé par le bras droit intact. Ce phénomène s'apparente fortement au mécanisme de BIR (*break induced replication*) observé entre deux loci ectopiques chez les eucaryotes (Pour revue : Anand *et al.*, 2013). La survenue de ces évènements de BIR est dépendante de la RH et pourrait être facilitée par la configuration spatiale du chromosome. En effet, il existe une interaction entre les protéines terminales fixées au niveau des télomères, permettant de maintenir une proximité physique entre les télomères (Tsai *et al.*, 2011), ce qui pourrait favoriser des évènements de RH non allélique entre les bras. Ce mécanisme a un impact important sur la plasticité du génome puisque dans le cas d'un évènement entre *hasL* et *hasR*, les 480 kb du bras gauche situés en amont d'*hasL* sont perdus tandis que les 840 kb situés en aval d'*hasR* sont dupliqués. Ainsi, la taille des régions terminales inversées passe de 200 à 840 kb (**Figure D5A**).

Lors de l'analyse des évènements de réparation, une différence majeure a été constatée entre les lignées dont l'évènement de remplacement de bras serait survenu entre *hasL* et *hasR* et les lignées dont l'évènement de remplacement de bras serait survenu entre σL et σR . En effet, pour *hasL* et *hasR*, les 6 lignées obtenues sont identiques et présentent un chromosome linéaire. En revanche, pour les 13 lignées dont l'évènement initial de RH se serait produit entre σL et σR , aucune ne présente un chromosome linéaire, un ou plusieurs évènements de RH et/ou RI entraînant une circularisation du chromosome s'étant produit à la suite du remplacement de bras. Afin d'essayer d'expliquer les conséquences différentes qu'entraîne un même évènement lorsqu'il survient via l'un ou l'autre des loci, les 90 kb situés entre *hasL* et σL (portion du chromosome perdue lors d'un remplacement du bras gauche par l'intermédiaire de $\sigma L/\sigma R$ mais toujours présente lors d'un remplacement du bras gauche par l'intermédiaire de *hasL/hasR*) ont été analysés afin d'identifier la nature des gènes présents. Les loci *SAM23877_0536* et *SAM23877_0537* qui correspondent respectivement aux gènes *tpg* et *tap* en font partie. Ainsi, après génération d'une DSB dans le bras gauche du chromosome, un évènement de RH ectopique entre σL et σR entraîne la

perte du locus *tpg-tap* et par conséquent abolit la protection des télomères par les protéines produites par ces gènes. Dans ce cas, l'évènement de remplacement de bras entrainerait probablement la formation d'un chromosome linéaire transitoire puisqu'en absence de protection par les protéines terminales, les extrémités chromosomiques vont rapidement être la cible de nucléases (Bao et Cohen, 2001, 2003). Ainsi, le maintien du chromosome sous forme linéaire n'est plus possible à long terme et seule une circularisation de ce dernier permet la survie de la bactérie (**figure D5B**). Dans nos expériences, la formation d'un chromosome linéaire transitoire pouvant subsister durant quelques générations pourrait se traduire par une population hétérogène, ce qui expliquerait la difficulté à obtenir des lignées stables et homogènes lors de nos étapes de sous-clonage. Par ailleurs, cette hypothèse permettrait également de comprendre pourquoi la proportion de chromosomes circulaires observés après génération d'une DSB à IKI-L (44/60) est plus élevée que celle observée dans les lignées obtenues après génération d'une DSB à IKI-R (7/40). En effet, si le remplacement du bras gauche par le bras droit après un évènement de BIR entre σ_L/σ_R entraîne la perte du locus *tap-tpg*, l'évènement réciproque entraîne en revanche une duplication de ce dernier.

Les éléments transposables : un substrat pour la RH intra-chromosomique

Pour 3 lignées analysées, la circularisation du chromosome résulte d'un évènement de RH ectopique entre différents couples de gènes dupliqués codant des transposases. Jusqu'à présent, la formation d'un chromosome circulaire décrite pour différentes espèces de *Streptomyces* a toujours été reliée à un évènement de recombinaison illégitime (Fischer *et al.*, 1997 Kameoka *et al.*, 1999 ; Inoue *et al.*, 2003) et ces résultats seraient donc les premiers exemples de circularisation spontanée du chromosome liée à la RH chez les *Streptomyces*. Aussi bien chez les eucaryotes que chez les procaryotes, les transposases dupliquées sont connues pour être des substrats de la RH (Raeside *et al.*, 2014). Afin d'évaluer l'impact qu'elles pourraient avoir sur la plasticité du génome, un inventaire des transposases présentes dans le génome de *S. ambofaciens* a été effectué. Au total 49 transposases réparties en 31 familles différentes ont été recensées. Parmi ces 31 familles, 9 comportent au moins deux copies de transposases qui partagent au minimum 99% d'identité nucléotidique. Sur ces 9 familles, 4 sont présentes dans les TIR ce qui explique la présence d'une copie de la transposase dans chaque bras. Les 5 autres familles comportent de 2 à 5 copies partielles ou complètes du locus, présentes majoritairement dans les bras

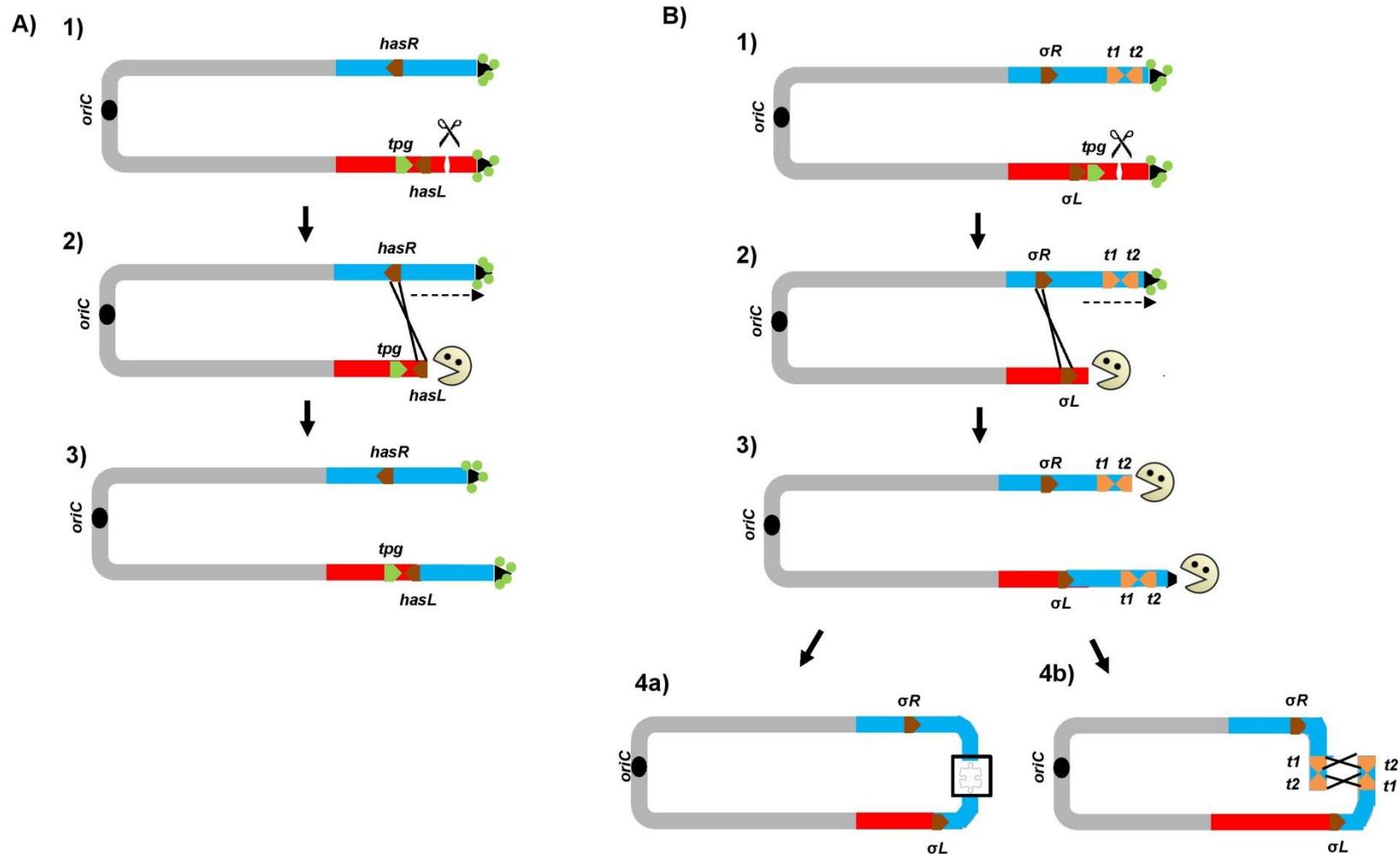


Figure D5 : Modèle de la prise en charge d'une DSB dans le bras gauche avec un évènement BIR-like initiateur de la réparation. **A)** Après génération de la cassure (ciseaux) (1), un évènement de BIR-like entre les deux gènes dupliqués *hasR* et *hasL* (marron) permet de recopier le bras droit (2) et d'obtenir ainsi un chromosome avec des TIR allant de *has* jusqu'aux télomères représentées par les flèches noires (3). Le gène *tpg* (vert) qui code les protéines terminales protégeant les télomères (cercles verts) étant situé en amont de *hasL*, la conformation linéaire du chromosome peut persister. **B)** Lorsque l'évènement de BIR-like a lieu entre les gènes dupliqués σ_L et σ_R (marron) (1-2), l'obtention d'un chromosome linéaire avec des TIR allant de σ aux télomères n'est que transitoire car le gène *tpg* (vert) est perdu et les télomères sont la cible de dégradation par les nucléases car elles ne sont plus protégées par les protéines terminales. (4). Ainsi, la survie de la bactérie est permise uniquement par la circularisation du chromosome soit par un évènement de RH entre gènes dupliqués (4A ; entre *t1* du bras droit et *t2* ou inversement (beige) (4b)) ou par un ou plusieurs évènements de RI possibles grâce à la libération des extrémités terminales (4a).

mais parfois également dans la région core et disposées en orientation directe ou opposée. En fonction de la localisation de DSB potentielles, ces transposases peuvent être source de RH ectopique et entraîner des réarrangements majeurs du chromosome.

Un cluster de gènes du métabolisme secondaire comme source de séquences répétées

Dans 5 lignées, la réparation de la cassure entraînerait une amplification en tandem à fort nombre de copies d'un *locus* situé au sein du cluster des polykétides synthases (PKS) de type I, responsable de la production des stambomycines (Laureti *et al.*, 2010). Young et Cullum (1987) puis Yanai *et al.* (2006) ont proposé un modèle fondé sur la réplication par cercle roulant pour expliquer la formation d'amplification dans des régions contenant des répétitions directes (**figure D6**).

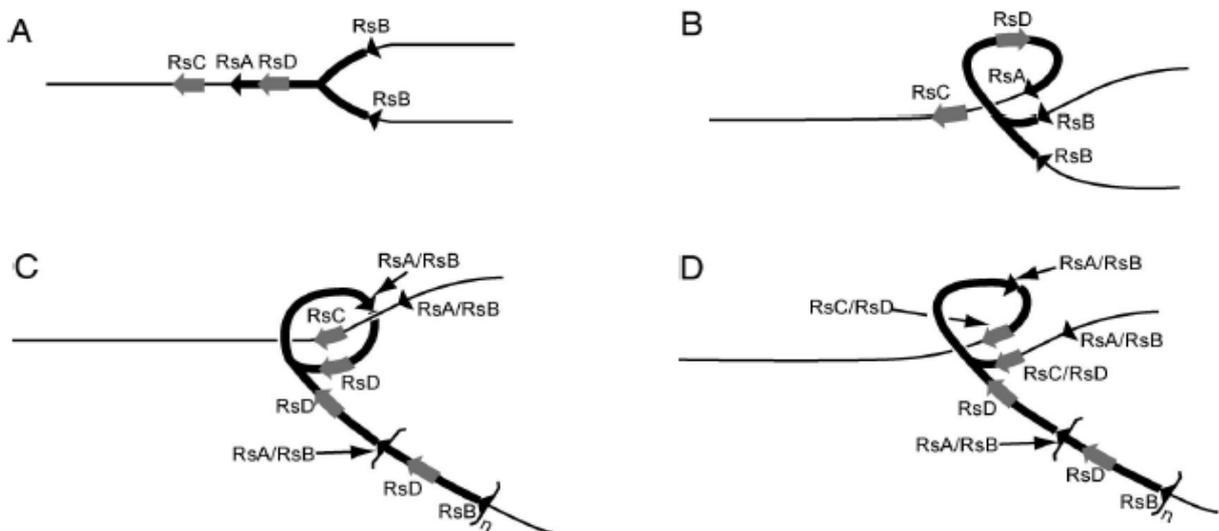


Figure D6 : Représentation schématique de l'amplification d'une AUD par réplication en cercle roulant (Yanai *et al.* 2006). Les différentes étapes sont détaillées dans le texte.

Durant le mécanisme de réplication, le piégeage d'une fourche de réplication résulte d'un évènement de RH entre deux séquences répétées positionnées dans la même orientation en amont et en aval du motif amplifié (représentées par RsA et RsB sur la **figure D6**), nommé unité amplifiable d'ADN (AUD). Le piège dans lequel est prise la machinerie répllicative entraîne la réplication intensive de l'AUD. La résolution de cette structure en cercle roulant se fait après un évènement de recombinaison entre deux autres séquences répétées (représentées par RsC et RsD sur la **figure D6**). Dans les génomes analysés, la taille du motif réitéré varie de 14 à 90 kb. L'analyse des régions amont et aval révèle

systématiquement la présence de séquences de plusieurs centaines de bases qui partagent plus de 90% d'homologie et qui sont probablement responsables de l'évènement d'amplification. Etant donné que ce phénomène est lié à la réplication, il est fort possible que l'amplification de ces AUD soit liée à l'évènement de BIR, ce qui expliquerait pourquoi la perte d'ADN est une caractéristique spécifique des évènements d'amplification chez les *Streptomyces* (Yanai *et al.*, 2006). Par ailleurs, pour l'ensemble de ces lignées, la présence des bornes de circularisation du chromosome au sein de l'AUD démontre l'importance du phénomène dans le processus de réparation des DSB au *locus* IKI-L. Cependant, il n'a pas été possible d'identifier les bornes exactes de la circularisation, et donc de conclure quant à l'implication de la RH ou de la RI.

1.2.2 La recombinaison illégitime a un impact majeur sur le génome

Bien que dans l'ensemble, le nombre d'évènements de RI soit plus faible que ceux de RH, les conséquences sur la structure du génome ne sont pas moins importantes. L'existence de mécanismes de RI est un outil formidable pour l'augmentation de la diversité génomique au sein d'une population. En effet, les mécanismes de RI n'ont pas la contrainte d'utilisation d'une matrice homologue, et la présence de micro-homologies ou simplement de deux extrémités franches permet la réparation d'une molécule d'ADN rompue. Ainsi, contrairement à la RH où les évènements sont contraints par la localisation et l'orientation des séquences répétées, la RI peut engendrer des structures chromosomiques très diverses.

1.3 Prise en charge des DSB chez d'autres microorganismes

Il est très difficile d'effectuer un comparatif entre les résultats obtenus chez *S. ambofaciens* et ceux décrits dans la littérature pour d'autres bactéries du fait de la linéarité et de l'organisation du chromosome des *Streptomyces*. Chez les rares bactéries qui possèdent un chromosome linéaire (*Borrelia*, *Agrobacterium*...) ce genre d'analyses n'a jusqu'à présent jamais été effectué. Récemment une expérience utilisant le système CRISPR-Cas9 a été mise en place chez *E. coli* afin de cibler des loci particuliers et d'analyser l'impact de l'induction des cassures sur le génome (Cui et Bikard, 2016). En absence d'une copie homologue intacte servant de matrice pour la RH, la majorité des cellules meurent par absence de réparation de la DSB. Les quelques survivants présentent de larges réarrangements résultants de recombinaison non allélique entre des séquences répétées REP ou d'évènements de recombinaison illégitime entraînant des délétions supérieures à 1 kb. L'intégration du système NHEJ de *Mycobacterium* chez *E. coli* n'augmente pas la capacité

de la bactérie à survivre aux DSB induites par Cas9. En revanche, en présence du mécanisme de NHEJ, la taille des délétions engendrées lors de la réparation est moins grande qu'en son absence (<300 pb).

Une analyse similaire à celle que nous avons mise en place pour *Streptomyces* a été effectuée chez la levure *Saccharomyces cerevisiae* (Ricchetti *et al.* 2003). L'utilisation de la méganucléase I-SceI a permis de cibler des loci particuliers et ainsi d'induire des DSB tout au long du chromosome XI de levures haploïdes pour lesquelles le remplacement allélique par RH n'est pas possible. Chez *S. cerevisiae* les régions subtélomériques ne contiennent que très peu de gènes, dont aucun n'est essentiel. L'analyse des évènements de réparation aux différents sites de cassure, révèle une utilisation quasi-exclusive du NHEJ dans la région centrale du chromosome entraînant des changements mineurs pour le génome. A mesure que l'on s'approche des extrémités du chromosome, le gradient d'utilisation du NHEJ diminue pour laisser place à d'autres mécanismes entraînant des réarrangements majeurs (Ricchetti *et al.* 2003). Parmi ces modifications majeures, il y a notamment des évènements BIR qui permettent de recopier le bras d'un autre chromosome par l'intermédiaire de recombinaison non allélique entre des gènes dupliqués. Par ailleurs, des fusions de chromosomes résultant d'évènements de RI indépendants du NHEJ et favorisés par la présence de micro-homologies ont également été observées (Ricchetti *et al.* 2003). Ainsi, de la même manière que chez *S. ambofaciens*, les conséquences qu'entraîne la réparation des DSB sont variables selon les régions du chromosome, la présence de gènes essentiels étant un frein majeur au maintien de réarrangements génomiques de grande ampleur.

2. Existence d'un mécanisme de recombinaison illégitime alternatif

2.1 Caractéristiques du mécanisme

L'analyse des séquençages de souches obtenues après induction de DSB ciblées dans le bras gauche révèle des évènements de réparation par RI même dans un contexte déficient pour le NHEJ (*ΔligCD*), ce qui suggère l'existence d'un mécanisme de RI alternatif. L'existence d'un mécanisme de RI indépendant du NHEJ a déjà été décrite chez différents organismes eucaryotes. Il est soit appelé MMEJ pour « *microhomology mediated end-joining* » ou encore A-EJ pour « *alternative end-joining* » (Pour revue : Deriano et Roth, 2013 ; Frit *et al.*, 2013). Il se caractérise également par une ligation médiée par des microhomologies de séquence entre les deux extrémités et est accompagné de délétions pouvant atteindre plusieurs kilobases.

Chez les procaryotes, et plus spécifiquement les bactéries, seuls deux exemples d'A-EJ ont été décrits. Chez *E. coli* qui est dépourvue de NHEJ classique, l'A-EJ semble être la seule alternative à la RH (Chayot *et al.*, 2009). Comme chez les eucaryotes, le système repose sur l'utilisation de microhomologies après une large résection des extrémités. En revanche, chez *Mycobacterium* l'A-EJ ne concernerait que la réparation des cassures à extrémités 3' sortantes et la ligation se ferait sans traitement des extrémités (Aniukwu *et al.*, 2008).

Chez *S. ambofaciens*, les cicatrices de réparation obtenues lors du séquençage des génomes réarrangés ne nous permettent pas de préciser la nature de la réparation dépendante de l'A-EJ. En effet, dans le contexte *AligCD*, unique contexte génétique où la déficience en NHEJ est avérée, seuls deux évènements de RI ont pu être observés ce qui ne permet pas de conclure. Cependant, dans les deux cas, les délétions associées à la recombinaison sont parmi les plus grandes observées, c'est à-dire 1,3 et 2,1 Mb. De plus, pour l'un des deux réarrangements, la ligation a été médiée par une microhomologie de 6 nucléotides, taille supérieure à la quasi-totalité des microhomologies observées dans les évènements NHEJ dépendants obtenus dans la région centrale du chromosome. Les caractéristiques de ces deux réarrangements suggèrent que l'A-EJ chez *Streptomyces* est semblable à celui des eucaryotes.

2.2 Acteurs possibles

L'A-EJ n'est en réalité pas considéré comme un mécanisme à part entière, c'est-à-dire mettant en œuvre des acteurs spécifiques, mais plutôt comme un système opportuniste composé d'acteurs impliqués dans différents mécanismes de réparation d'ADN (**Figure D7**). Chez les eucaryotes, il se définit comme l'ensemble des évènements de RI indépendants de Ku70/80 et LigIV. L'intervention de différents acteurs peut être disséquée en trois étapes : (i) reconnaissance de la cassure et fixation, (ii) traitement des extrémités (iii) et ligation.

Bien que son rôle reste peu défini, la protéine PARP-1 interviendrait dans le mécanisme de réparation par excision de bases (BER ; pour revue : Krokan et Bjoras ,2013). Après glycosylation d'une base endommagée, le clivage du site abasique engendre la formation d'une cassure simple-brin qui serait utilisée comme substrat pour PARP-1 (De Vos *et al.*, 2012). De la même manière, PARP-1 est également un senseur des DSB, ce qui en fait l'acteur principal de l'étape initiale de l'A-EJ (Frit *et al.*, 2013). Après fixation au niveau de la cassure, PARP-1 sert de plateforme pour le recrutement direct ou indirect de l'ensemble des composants qui interviennent dans le traitement des extrémités. Parmi eux, le complexe MRN chez les mammifères (Mre11, Rad50 et Nbs1) ou MRX chez la levure

(Mre11, Rad50 et Xrs2) qui joue un rôle central dans la recombinaison homologue, permet le pontage des extrémités. Il intervient également en collaboration avec la protéine Clt lors de l'étape de résection des extrémités. Enfin, la ligase III stabilisée par XRCC1 permet de liguer les extrémités par l'intermédiaire de micro-homologies. En absence de la ligase III, un effet back-up peut être assuré grâce à la ligase I (Simsek *et al.*, 2011). L'A-EJ et l'intervention de ses acteurs principaux sont résumés dans la **figure D8**.

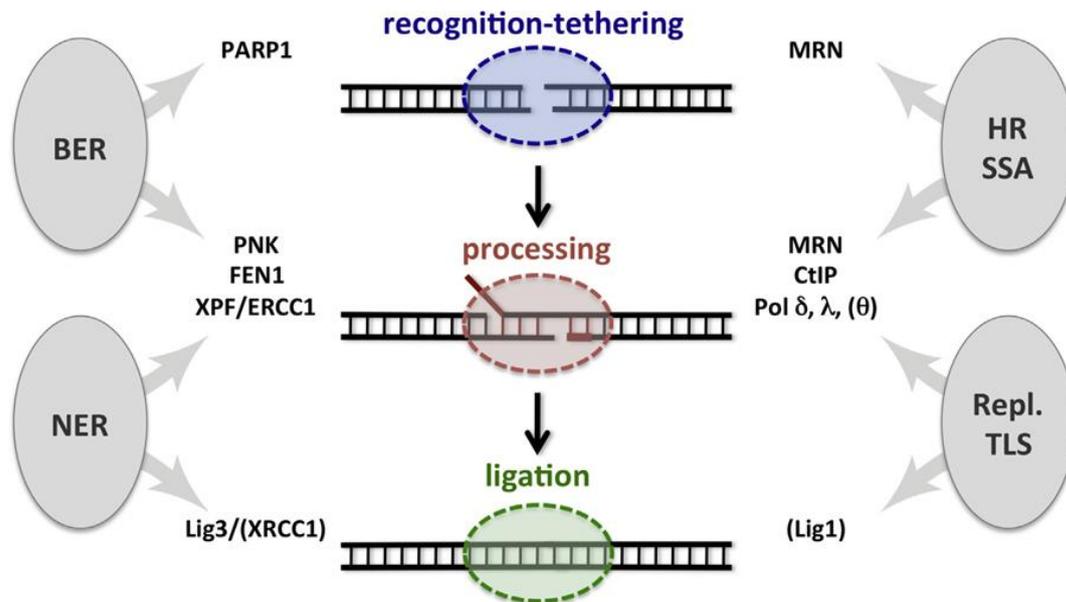


Figure D7 : Etapes de l'A-EJ eucaryote ainsi que des acteurs qui y sont associés. De chaque côté de la figure sont indiqués les mécanismes dans lesquels les acteurs de l'A-EJ interviennent également. La signification des acronymes est la suivante : BER = Base excision repair ; NER = Nucleotide excision repair ; HR = Homologous recombination ; SSA = Single strand annealing ; Repl. = Replication ; TLS = Translesion synthesis (Frit *et al.*, 2013)

Les acteurs de l'A-EJ bactérien restent hypothétiques. Bien que le mécanisme de BER existe chez les bactéries, aucun homologue de PARP-1 n'a été identifié à ce jour. Chez *E. coli*, il a été montré que l'étape de résection des extrémités lors de l'A-EJ est réalisée par le complexe hélicase-nucléase RecBCD (Chayot *et al.*, 2009). La ligation des extrémités de la cassure se fait grâce à la ligase A, une protéine essentielle qui intervient dans la ligation des fragments d'Okasaki lors de la réplication (Chayot *et al.*, 2009). Dans le mécanisme d'A-EJ décrit chez *Mycobacterium*, la ligation semble également dépendante de LigA (Aniukwu *et al.*, 2008). Chez *Streptomyces*, RecBCD n'existe pas et il est envisageable que la résection des extrémités lors de l'A-EJ soit prise en charge par AdnAB, son homologue fonctionnel. Enfin, en absence de LigC et LigD, l'intervention de LigA apparaît comme un acteur probable pour la ligation des extrémités lors de l'A-EJ.

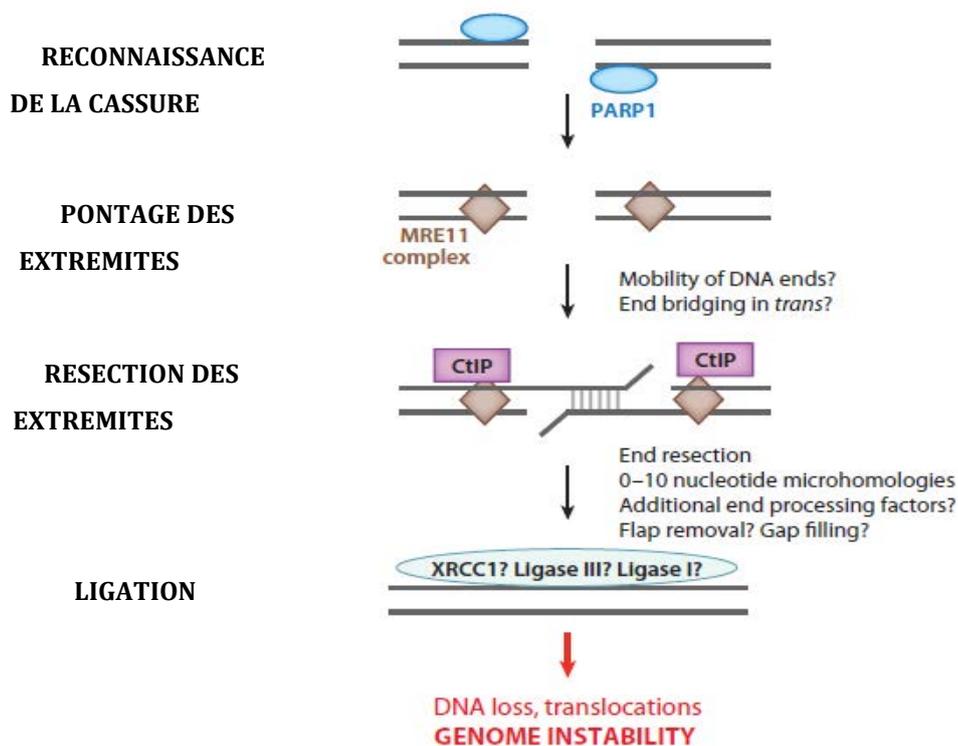


Figure D8 : Représentation schématique de l'A-EJ chez les eucaryotes (D'après Deriano et Roth, 2013)

3. Régulation de l'intervention RH/NHEJ

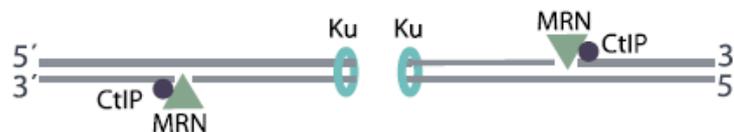
Une compétition entre les acteurs du RH et du NHEJ pour la prise en charge des DSB chez les eucaryotes

Chez les eucaryotes, l'intervention préférentielle de la RH ou du NHEJ est fortement dépendante du cycle cellulaire. En effet, chez la levure, l'expression de *ku70* augmente avant la phase G1, durant laquelle l'absence de matrice homologue empêche la RH de jouer son rôle (Clerici *et al.*, 2008). Le choix du mécanisme pour la prise en charge des DSB semble se faire au moment de la résection des extrémités de la cassure. Une délétion des gènes *ku* entraîne une augmentation de la résection des extrémités par le complexe exonucléasique initiant la RH (Lee *et al.*, 1998), tandis que l'augmentation de la prise en charge des DSB par le NHEJ ne se fait que lorsque les extrémités de la DSB n'ont pas encore été prises en charge par les enzymes pré-synaptiques de la RH (Stark *et al.*, 2004).

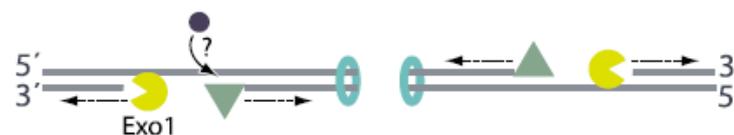
Des études effectuées chez *S. cerevisiae* tendent à montrer qu'un élément central de la régulation de la prise en charge des DSB par les mécanismes de réparation serait le complexe MRX (Mre11-Rad50-XRS) équivalent de MRN chez les mammifères (Mre11-Rad50-Nbs1). Il a été décrit dans la partie précédente pour son rôle dans l'A-EJ mais il

intervient également dans la RH et le NHEJ. Dans la RH, le complexe MRX/N jouerait un rôle d'initiateur du mécanisme et interviendrait dans le pontage des extrémités. Dans le NHEJ, il permettrait également le pontage des extrémités et le recrutement de certains facteurs du NHEJ (pour revue : Daley *et al.*, 2005). Lors de la formation d'une cassure, le recrutement de Ku et MRX/N se fait quasi-simultanément mais de manière indépendante (Wu *et al.*, 2008) et de façon plus rapide que celui des acteurs de l'étape pré-synaptique de la RH. La fixation de Ku aux extrémités de la cassure protège ces dernières d'une dégradation par le complexe exonucléasique de la RH et inhibe ainsi l'action de la RH (Zhang *et al.*, 2007). Le complexe MRX/N contrôle la protection des extrémités par Ku et peut aussi bien favoriser sa stabilisation que sa dissociation (Zhang *et al.* 2007). La balance se ferait grâce à l'interaction du complexe CtlP (constitué des protéines Ctl1 et Sae2) avec MRX/N après phosphorylation de CtlP par les protéines CDK (pour cyclin-dependent kinases) lors d'un processus de régulation dépendant du cycle cellulaire (Pour revue : Chapman *et al.*, 2012). Le mécanisme de régulation de MRX/N CtlP reste à ce jour inconnu. Le modèle proposé par l'équipe de Chapman sur l'initiation de la résection des extrémités de l'ADN par MRX/N est présenté dans la **figure D9** (Chapman *et al.*, 2012).

1) L'activité de coupure simple-brin de l'endonucléase Mre11 initie la résection en amont de la DSB sur laquelle est fixé Ku.



2) La résection bidirectionnelle est dépendante à la fois de Mre11 et Exo1



3) La dégradation de 3' vers 5' favorise le décrochage de Ku

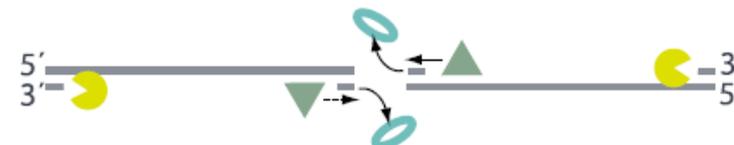


Figure D9 : Modèle eucaryote de l'initiation de la résection des extrémités d'une DSB dépendante des complexes MRN et CtlP (Chapman *et al.*, 2012)

Après activation, le complexe CtlP-MRX/N initie la résection des extrémités environ 300 pb en amont de la cassure grâce à l'activité endonucléasique de Mre11 (**figure D9-1**). Le

recrutement d'ExoI permet la mise en place d'une activité exonucléasique bidirectionnelle (**figure D9-2**). La dégradation de 5' vers 3' facilite le décrochage de l'hétérodimère Ku et permet la poursuite de la réparation par RH (**figure D9-3**). La robustesse de ce modèle est renforcée par des études effectuées chez la levure qui montrent que (i) la protection des extrémités de la cassure est plus efficace en absence de MRX et Sae2 et (ii) que l'initiation de la RH nécessite uniquement ExoI lorsque Ku est absent (Mimitou et Symington, 2010 ; Mathiasen *et al.*, 2014).

Des informations parcellaires suggérant une compétition NHEJ/RH chez les procaryotes

Chez *Mycobacterium*, des analyses *in vitro* ont révélé que Ku empêche la dégradation d'extrémités d'ADN par AdnAB (Sinha *et al.*, 2009). De plus, une augmentation de l'efficacité de la RH est observée lorsque le NHEJ est muté par délétion de *ku* mais pas par délétion de *ligD* (Gupta *et al.*, 2011). Ces résultats qui concordent avec les observations faites chez *S. cerevisiae* (Zhang *et al.*, 2007) suggèrent que Ku et AdnAB sont en compétition pour la prise en charge d'extrémités d'ADN libres et que la forte affinité de Ku pour les extrémités d'ADN bloque l'intervention d'AdnAB. Les résultats obtenus chez *Mycobacterium* sont cependant unidirectionnels, puisque la délétion de *recA* n'augmente pas l'efficacité de prise en charge des DSB par le NHEJ (Gupta *et al.*, 2011).

Une étude effectuée chez *S. avermitilis* a permis de constater que la délétion des deux gènes *ku* engendre une augmentation des fréquences de mutagenèse ciblée par RH (Zhang *et al.*, 2012), ce qui concorde avec les observations faites pour *Mycobacterium* (Gupta *et al.*, 2011). Bien que notre système expérimental d'induction des DSB ne nous permette pas en l'état de mesurer la prise en charge d'un DSB par les différents mécanismes de réparation, l'analyse des taux de survie des différents contextes génétiques (*recA*- : [RH-], Δ *ligCD* et Δ *kuABC* : [NHEJ-] et leurs combinaisons [RH- NHEJ-]) aux rayons ionisants (1 kGy) est riche d'enseignement. Ainsi, un mutant Δ *kuABC* et Δ *ligCD* [NHEJ-] est 7 fois plus sensible que la souche sauvage [RH+ NHEJ+], un mutant Δ *recA* ([RH-] est 50 fois plus sensible que la souche sauvage tandis que les mutants Δ *kuABC* Δ *recA* et Δ *ligCD* Δ *recA* ([RH-NHEJ-]) sont 10 000 fois plus sensibles que la souche sauvage. Ainsi l'impact d'une délétion des gènes *ku* serait plus importante lorsque la RH est déficiente ce qui suggère qu'en absence des gènes *ku*, la RH compense l'absence du NHEJ (**figure D10A**). Réciproquement, les résultats suggèrent que le NHEJ est accru en absence de RH (**figure D10B**). Il serait intéressant de tester cette hypothèse par exemple en mesurant le taux d'expression des gènes *ku* et *lig* dans un contexte muté pour la RH et de le comparer à l'expression de ces mêmes gènes dans une

souche sauvage, si toutefois cette augmentation d'activité est liée à une régulation transcriptionnelle ou post-transcriptionnelle.

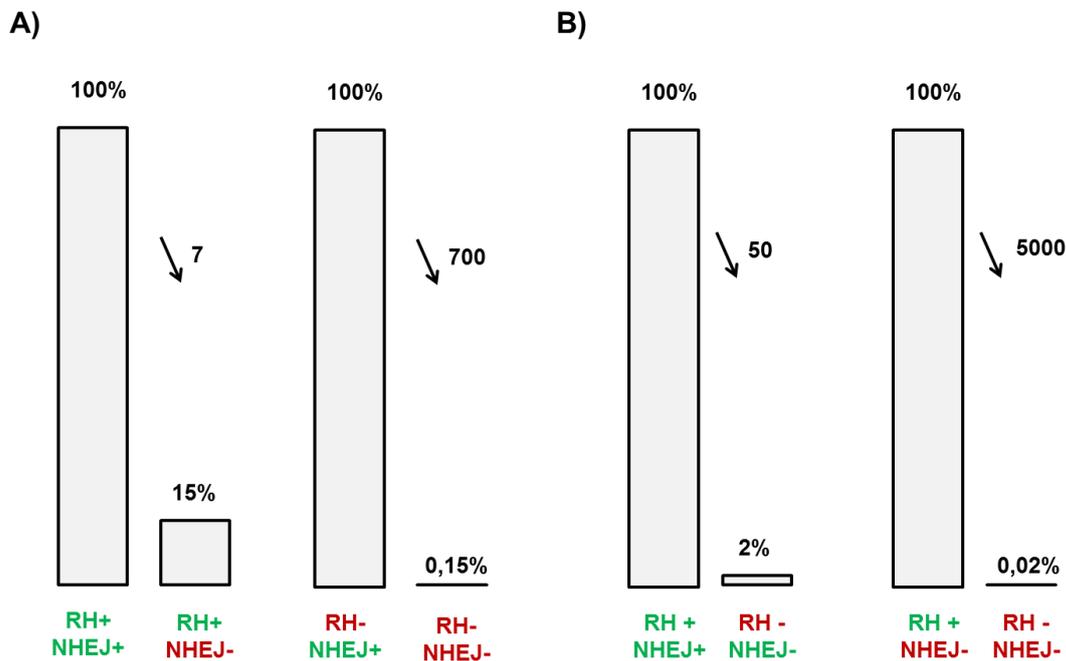


Figure D10 : Taux de survie après exposition aux rayons électroniques exprimée relativement à un contexte génétique donné. A) Impact de la délétion du NHEJ sur le taux de survie des spores après exposition aux rayons électroniques par rapport à un contexte WT (colonnes de gauche) ou un contexte RH- (colonnes de droite) B) Impact de la délétion de la RH sur le taux de survie des spores après exposition aux rayons électroniques par rapport à un contexte WT (colonnes de gauche) ou un contexte NHEJ- (colonnes de droite). La flèche vers le bas indique le facteur de chute entre les deux contextes

Il existe peu d'information sur la régulation des gènes du NHEJ chez les bactéries. Chez *B. subtilis*, cependant, ces derniers sont sous le contrôle transcriptionnel de σ^G , facteur sigma qui intervient tardivement dans la formation de l'endospore (Wang *et al.*, 2006). Chez *Mycobacterium*, la découverte d'une interaction entre UvrD et Ku ainsi que l'intervention d'UvrD dans d'autres mécanismes de réparation comme l'excision de bases et la RH suggère qu'UvrD pourrait jouer un rôle d'intermédiaire dans l'interaction entre les différents mécanismes de réparation de l'ADN (Sinha *et al.*, 2007). Chez *S. ambifaciens* où deux allèles UvrD sont présents il serait intéressant de tester leur rôle dans la régulation croisée de la RH et de la RI en générant des mutants simples et doubles de $\Delta uvrD1$ et $\Delta uvrD2$. De plus, des tests d'interaction entre les protéines UvrD et les différents acteurs du NHEJ et de la RH pourraient également apporter des réponses sur leur implication. Plus récemment,

l'implication dans le NHEJ d'un gène qui code une protéine de la famille Sir2 a été mise en évidence chez *Mycobacterium* (Li *et al.*, 2011). La protéine Sir2 qui est une déacétylase NAD dépendante interagit physiquement avec Ku et LigD et sa délétion diminue la fréquence de réparation par NHEJ (Li *et al.*, 2011). Des analyses complémentaires ont permis de mettre en évidence une acétylation de Ku qui permettrait une régulation de la quantité de cette dernière au sein de la cellule (Zhou *et al.*, 2015). *S. ambifaciens* possède un homologue de *sir2* (SAM23877_6553) dont le rôle dans le NHEJ pourrait également être testé.

PERSPECTIVES

Les résultats acquis sur les mécanismes de recombinaison chez les *Streptomyces* ont permis d'établir un lien entre réparation des DSB et plasticité génomique. En effet, l'induction puis la réparation de DSB à l'intérieur des bras chromosomiques de *Streptomyces ambofaciens* génère l'ensemble des réarrangements génomiques observés chez des mutants spontanés obtenus en conditions de laboratoire, voire de souches issues de l'environnement. La réparation de DSB permet donc d'expliquer la plasticité génomique des *Streptomyces* observée à maintes reprises. De plus, ces travaux ont montré que le mécanisme de NHEJ pouvait permettre l'insertion de matériel génétique au point de réparation ouvrant la possibilité d'intégration d'information génétique issu du transfert horizontal. Ainsi, ces travaux ouvrent des pistes de recherche qui permettront d'aborder les mécanismes évolutifs des *Streptomyces*. Comment à partir d'un génome ancestral unique et linéaire (hypothèse la plus parcimonieuse à l'extrême) est-on parvenu à la diversité génétique extraordinaire des génomes contemporains ? Comment expliquer la compartimentation génétique du chromosome, c'est-à-dire la présence d'une région centrale conservée au travers des espèces, et des bras chromosomiques variables ? Enfin, comment expliquer le phénomène de synténie dégénérée (**figure D11**), c'est-à-dire la diminution progressive de la conservation de l'ordre des gènes ancestraux du cœur du chromosome vers les extrémités ?

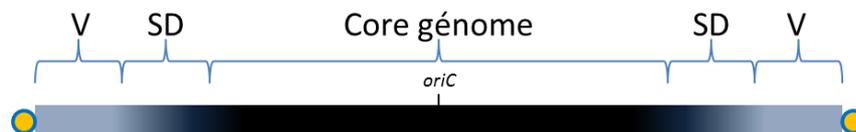


Figure D11 : Représentation schématique de la compartimentation du chromosome des *Streptomyces* révélée par la comparaison génomique. En noir profond, la région centrale synténique lorsque les génomes de toutes les espèces séquencées sont comparées deux à deux. En dégradé de gris, la région de synténie dégénérée (SD) et en gris pâle, la région variable.

La synténie dégénérée pose la question de la régio-spécificité des événements d'insertion et délétion (indel) le long du génome. En effet, on observe dans ces régions une accumulation d'indels (de 1 à 10 gènes) gommant progressivement la synténie ancestrale (Choulet *et al.*, 2006). Si les insertions et délétions résultent, au moins pour part, de la réparation de DSB, alors on peut émettre l'hypothèse que soit les cassures augmentent en fréquence le long du chromosome en direction des extrémités, soit que la prise en charge

des DSB le long du génome est variable. Ainsi, il serait intéressant de tester cette hypothèse en observant la prise en charge d'une DSB placée en différentes positions dans les régions terminales. Cependant, comme l'ont démontré ces travaux, l'induction d'une cassure artificielle ne permet pas d'accéder à la réparation du site souhaité puisqu'une forte activité nucléasique engendre des grandes délétions. Comment aborder cette question ? Une façon de procéder serait de tester la vitesse d'évolution du chromosome par l'intermédiaire d'expériences d'évolution expérimentale sur le long terme à l'image des lignées d'*Escherichia coli* mises en place par Richard Lenski à l'Université du Michigan (USA) et Dominique Schneider à l'Université Joseph Fourier (Grenoble). Pour cela, nous pourrions utiliser les différents contextes construits (déficient pour la RI ou la RH) pour observer les différents parcours évolutifs qu'emprunterait le génome des *Streptomyces* sur le long terme. Une expérience de co-culture entre *S. ambofaciens* et une espèce proche *S. coelicolor*, par exemple, pourrait également permettre de tester l'impact de la RI sur l'intégration de matériel exogène (*S. ambofaciens* et *S. coelicolor* sont susceptibles d'échanger du matériel génétique par conjugaison). Compte tenu de la vitesse de diversification du chromosome des *Streptomyces* observée dans des populations clonales de la rhizosphère (travaux de Maxime Toussaint en thèse à Dynamic, Dir. C. Bontemps, P. Leblond), il est raisonnable de penser qu'une année d'expérience (soit environ 1 300 générations ; 52 cycles de sporulation chacun correspondant à 25 divisions cellulaires) serait suffisante pour observer des schémas évolutifs.

Des expériences similaires permettraient, en complément de la reconstitution *in vitro* de la machinerie enzymatique du NHEJ (collaboration avec S. McGovern et F. Lecointe, MICALIS, INRA) de tester le rôle des gènes NHEJ-like variables dans l'évolution chromosomique. En effet, une question soulevée par ces travaux est bien l'impact de ces gènes (acquis par transfert ou 'simplement' conservés dans certains isolats), sur la capacité recombinatoire du génome d'une souche. Y a-t-il une diversité de la compétence à recombiner au sein d'une population naturelle ? La variabilité de la recombinaison est-elle un moteur d'évolution d'une population dans son écosystème ?

Références bibliographiques

- Adams, D. E., and West, S. C. (1996). Bypass of DNA heterologies during RuvAB-mediated three-and four-strand branch migration. *J. Mol. Biol.* 263, 582–596.
- Ali, N., Herron, P. R., Evans, M. C., and Dyson, P. J. (2002). Osmotic regulation of the *Streptomyces lividans* thiostrepton-inducible promoter, *ptipA*. *Microbiol.* 148, 381–390.
- Alt, F.W., Oltz, E. M., Young, F., Gorman, J., Taccioli, G., and Chen, J. (1992). VDJ recombination. *Immunol. Today* 13, 306-314.
- Alvarez-Martinez, C. E., and Christie, P. J. (2009). Biological Diversity of Prokaryotic Type IV Secretion Systems. *Microbiol. Mol. Biol. Rev.* 73, 775–808. doi:10.1128/MMBR.00023-09.
- Amundsen, S. K., Neiman, A. M., Thibodeaux, S. M., and Smith, G. R. (1990). Genetic dissection of the biochemical activities of RecBCD enzyme. *Genet.* 126, 25–40.
- Anand, R. P., Lovett, S. T., and Haber, J. E. (2013). Break-Induced DNA Replication. *Cold Spring Harb. Persp. Biol.* 5, a010397–a010397. doi:10.1101/cshperspect.a010397.
- Aniukwu, J., Glickman, M. S., and Shuman, S. (2008). The pathways and outcomes of mycobacterial NHEJ depend on the structure of the broken DNA ends. *Genes Dev.* 22, 512–527. doi:10.1101/gad.1631908.
- Aravind, L. and Koonin, E. V. (2001). Prokaryotic Homologs of the Eukaryotic DNA-End-Binding Protein Ku, Novel Domains in the Ku Protein and Prediction of a Prokaryotic Double-Strand Break Repair System. *Genome Res.* 11, 1365–1374. doi:10.1101/gr.181001.
- Aravind, L., Makarova K. S., and Koonin, E. V. (2000). Holliday junction resolvases and related nucleases: identification of new families, phyletic distribution and evolutionary trajectories. *Nucleic Acids Res.* 28, 3417–3432.
- Ayora, S., Carrasco, B., Cárdenas, P. P., César, C. E., Cañas, C., Yadav, T., *et al.* (2011). Double-strand break repair in bacteria: a view from *Bacillus subtilis*. *FEMS Microbiol. Rev.* 35, 1055–1081. doi:10.1111/j.1574-6976.2011.00272.x.
- Ayora, S., Carrasco, B., Doncel, E., Lurz, R., and Alonso, J. C. (2004). *Bacillus subtilis* RecU protein cleaves Holliday junctions and anneals single-stranded DNA. *Proc. Natl. Acad. Sci. U. S. A.* 101, 452–457.
- Azeroglu, B., Mawer, J. S. P., Cockram, C. A., White, M. A., Hasan, A. M. M., Filatenkova, M., *et al.* (2016). RecG Directs DNA Synthesis during Double-Strand Break Repair. *PLOS Genet.* 12, e1005799. doi:10.1371/journal.pgen.1005799.
- Babu, M., Beloglazova, N., Flick, R., Graham, C., Skarina, T., Nocek, B., *et al.* (2011). A dual function of the CRISPR-Cas system in bacterial antiviral immunity and DNA repair: Branched DNA nuclease YgbT. *Mol. Microbiol.* 79, 484–502. doi:10.1111/j.1365-2958.2010.07465.x.
- Bao, K., and Cohen, S. N. (2001). Terminal proteins essential for the replication of linear plasmids and chromosomes in *Streptomyces*. *Genes Dev.* 15, 1518–1527.

- Bao, K., and Cohen, S. N. (2003). Recruitment of terminal protein to the ends of *Streptomyces* linear plasmids and chromosomes by a novel telomere-binding protein essential for linear DNA replication. *Genes Dev.* 17, 774-785.
- Bartlett, E. J., Brissett, N. C., and Doherty, A. J. (2013). Ribonucleolytic resection is required for repair of strand displaced nonhomologous end-joining intermediates. *Proc. Natl. Acad. Sci. U. S. A.* 110, E1984-E1991. doi:10.1073/pnas.1302616110.
- Bartlett, E. J., Brissett, N. C., Plocinski, P., Carlberg, T., and Doherty, A. J. (2016). Molecular basis for DNA strand displacement by NHEJ repair polymerases. *Nucleic Acids Res.* 44, 2173-2186. doi:10.1093/nar/gkv965.
- Beam, C. E., Saveson, C. J., and Lovett, S. T. (2002). Role for *radA/sms* in Recombination Intermediate Processing in *Escherichia coli*. *J. Bacteriol.* 184, 6836-6844. doi:10.1128/JB.184.24.6836-6844.2002.
- Bentley, S. D., Chater, K. F., Cerdeno-Tarraga, A.-M., Challis, G. L., Thomson, N. R., James, K. D., et al. (2002). Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3 (2). *Nature* 417, 141-147.
- Bhattacharai, H., Gupta, R., and Glickman, M. S. (2014). DNA Ligase C1 Mediates the LigD-Independent Nonhomologous End-Joining Pathway of *Mycobacterium smegmatis*. *J. Bacteriol.* 196, 3366-3376. doi:10.1128/JB.01832-14.
- Birch, A., Häusler, A., Rüttener, C., and Hütter, R. (1991). Chromosomal deletion and rearrangement in *Streptomyces glaucescens*. *J. Bacteriol.* 173, 3531-3538.
- Birch, A., Häusler, A., Vögli, M., Krek, W., and Hütter, R. (1989). Extremely large chromosomal deletions are intimately involved in genetic instability and genomic rearrangements in *Streptomyces glaucescens*. *Mol. Gen. Genet.* 217, 447-458.
- Bowater, R., and Doherty, A. J. (2006). Making Ends Meet: Repairing Breaks in Bacterial DNA by Non-Homologous End-Joining. *PLoS Genet.* 2, e8. doi:10.1371/journal.pgen.0020008.
- Brissett, N. C., and Doherty, A. J. (2009). Repairing DNA double-strand breaks by the prokaryotic non-homologous end-joining pathway. *Biochem. Soc. Trans.* 37, 539-545. doi:10.1042/BST0370539.
- Burke, J., Schneider, D., and Westpheling, J. (2001). Generalized transduction in *Streptomyces coelicolor*. *Proc. Natl. Acad. Sci. U. S. A.* 98, 6289-6294.
- Carter, A. S., Tahmaseb, K., Compton, S. A., and Matson, S. W. (2012). Resolving Holliday Junctions with *Escherichia coli* UvrD Helicase. *J. Biol. Chem.* 287, 8126-8134. doi:10.1074/jbc.M111.314047.
- Chapman, J. R., Taylor, M. R. G., and Boulton, S. J. (2012). Playing the End Game: DNA Double-Strand Break Repair Pathway Choice. *Mol. Cell* 47, 497-510. doi:10.1016/j.molcel.2012.07.029.
- Chater, K. F. (1993). Genetics of differentiation in *Streptomyces*. *Annu. Rev. Microbiol.* 47,

685–711.

Chayot, R., Montagne, B., Mazel, D., and Ricchetti, M. (2010). An end-joining repair mechanism in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* 107, 2141–2146. doi:10.1073/pnas.0906355107.

Chédin, F., Ehrlich, S. D., and Kowalczykowski, S. C. (2000). The *Bacillus subtilis* AddAB helicase/nuclease is regulated by its cognate Chi sequence in vitro. *J. Mol. Biol.* 298, 7–20. doi:10.1006/jmbi.2000.3556.

Chédin, F., Handa, N., Dillingham, M. S., and Kowalczykowski, S. C. (2006). The AddAB Helicase/Nuclease Forms a Stable Complex with Its Cognate Sequence During Translocation. *J. Biol. Chem.* 281, 18610–18617. doi:10.1074/jbc.M600882200.

Chédin, F., Noiroot, P., Biaudet, V. and Ehrlich, S. D (1998). A five-nucleotide sequence protects DNA from exonucleolytic degradation by AddAB, the RecBCD analogue of *Bacillus subtilis*. *Mol. Microbiol.* 29, 1369-1377.

Chen, C. W. (1996). Complications and implications of linear bacterial chromosomes. *Trends Genet.* 12, 192–196.

Chen, C. W., Huang, C.-H., Lee, H.-H., Tsai, H.-H., and Kirby, R. (2002). Once the circle has been broken: dynamics and evolution of *Streptomyces* chromosomes. *Trends Genet.* 18, 522–529.

Chiu, M. L., Folcher, M., Katoh, T., Puglia, A. M., Vohradsky, J., Yun, B.-S., et al. (1999). Broad Spectrum Thiopeptide Recognition Specificity of the *Streptomyces lividans* TipAL Protein and Its Role in Regulating Gene Expression. *J. Biol. Chem.* 274, 20578–20586.

Choulet, F., Aigle, B., Gallois, A., Mangenot, S., Gerbaud, C., Truong, C., et al. (2006). Evolution of the Terminal Regions of the *Streptomyces* Linear Chromosome. *Mol. Biol. Evol.* 23, 2361–2369. doi:10.1093/molbev/msl108.

Christie, P. J., Whitaker, N., and González-Rivera, C. (2014). Mechanism and structure of the bacterial type IV secretion systems. *Biochim. Biophys. Acta.* 1843, 1578–1591. doi:10.1016/j.bbamcr.2013.12.019.

Clerici, M., Mantiero, D., Guerini, I., Lucchini, G., and Longhese, M. P. (2008). The Yku70–Yku80 complex contributes to regulate double-strand break processing and checkpoint activation during the cell cycle. *EMBO Rep.* 9, 810–818. doi:10.1038/embor.2008.121.

Connelly, J. C., de Leau, E. S., Okely, E. A., and Leach, D. R. (1997). Overexpression, purification, and characterization of the SbcCD protein from *Escherichia coli*. *J. Biol. Chem.* 272, 19819–19826.

Connolly, B., Parsons, C. A., Benson, F. E., Dunderdale, H. J., Sharples, G. J., Lloyd, R. G., et al. (1991). Resolution of Holliday junctions in vitro requires the *Escherichia coli* *ruvC* gene product. *Proc. Natl. Acad. Sci. U. S. A.* 88, 6063–6067.

Cooper, D. L., Boyle, D. C., and Lovett, S. T. (2015). Genetic analysis of *Escherichia coli* RadA: functional motifs and genetic interactions. *Mol. Microbiol.* 95, 769–779.

doi:10.1111/mmi.12899.

Cordero, O. X., and Polz, M. F. (2014). Explaining microbial genomic diversity in light of evolutionary ecology. *Nature Rev. Microbiol.* 12, 263–273. doi:10.1038/nrmicro3218.

Costes, A., and Lambert, S. (2012). Homologous Recombination as a Replication Fork Escort: Fork-Protection and Recovery. *Biomolecules.* 3, 39–71. doi:10.3390/biom3010039.

Cox, M. M. (2007). Motoring along with the bacterial RecA protein. *Nat. Rev. Mol. Cell Biol.* 8, 127–138. doi:10.1038/nrm2099.

Cui, L., and Bikard, D. (2016). Consequences of Cas9 cleavage in the chromosome of *Escherichia coli*. *Nucleic Acids Res.* 44, 4243–4251. doi:10.1093/nar/gkw223.

Daley, J. M., Palmbo, P. L., Wu, D., and Wilson, T. E. (2005). Nonhomologous end joining in yeast. *Annu. Rev. Genet.* 39, 431–451.

Davis, E. O., Springer, B., Gopaul, K. K., Papavinasasundaram, K. G., Sander, P., and Böttger, E. C. (2002). DNA damage induction of *recA* in *Mycobacterium tuberculosis* independently of RecA and LexA. *Mol. Microbiol.* 46, 791–800.

De Vos, M., Schreiber, V., and Dantzer, F. (2012). The diverse roles and clinical relevance of PARPs in DNA damage repair: Current state of the art. *Biochem. Pharmacol.* 84, 137–146. doi:10.1016/j.bcp.2012.03.018.

Decottignies, A. (2005). Capture of Extranuclear DNA at Fission Yeast Double-Strand Breaks. *Genet.* 171, 1535–1548. doi:10.1534/genetics.105.046144.

Della, M., Palmbo, P. L., Tseng H., Tonkin L. M., Daley J. M., Topper L. M. *et al.* (2004). Mycobacterial Ku and ligase proteins constitute a two-component NHEJ repair machine. *Science* 306, 683–685. doi:10.1126/science.1099824.

Demuyter, P., Leblond, P., Decaris, B., and Simonet, J.-M. (1988). Characterization of two families of spontaneously amplifiable units of DNA in *Streptomyces ambofaciens*. *Microbiol.* 134, 2001–2007.

Deriano, L., and Roth, D. B. (2013). Modernizing the Nonhomologous End-Joining Repertoire: Alternative and Classical NHEJ Share the Stage. *Annu. Rev. Genet.* 47, 433–455. doi:10.1146/annurev-genet-110711-155540.

Dillingham, M. S., and Kowalczykowski, S. C. (2008). RecBCD Enzyme and the Repair of Double-Stranded DNA Breaks. *Microbiol. Mol. Biol. Rev.* 72, 642–671. doi:10.1128/MMBR.00020-08.

Doherty, A. J., Jackson, S. P., and Weller, G. R. (2001). Identification of bacterial homologues of the Ku DNA repair proteins. *FEBS Lett.* 500, 186–188.

Dudášová, Z., Dudáš, A., and Chovanec, M. (2004). Non-homologous end-joining factors of *Saccharomyces cerevisiae*. *FEMS Microbiol. Rev.* 28, 581–601. doi:10.1016/j.femsre.2004.06.001.

- Dunderdale, H. J., Benzon, F. E., Parsons, C. A., Sharples, G. J., Lloyd, R. G., and West, S. C. (1991). Formation and resolution of recombination intermediates by *E. coli* RecA and RuvC proteins. *Nature* 354, 506-510.
- Fell, V. L., and Schild-Poulter, C. (2015). The Ku heterodimer: Function in DNA repair and beyond. *Mut. Res. Rev.* 763, 15–29. doi:10.1016/j.mrrev.2014.06.002.
- Fernández-Martínez, L. T., and Bibb, M. J. (2014). Use of the Meganuclease I-SceI of *Saccharomyces cerevisiae* to select for gene deletions in actinomycetes. *Sci. Rep.* 4, 7100. doi:10.1038/srep07100.
- Fischer, G., Decaris, B., and Leblond, P. (1997). Occurrence of deletions, associated with genetic instability in *Streptomyces ambofaciens*, is independent of the linearity of the chromosomal DNA. *J. Bacteriol.* 179, 4553–4558.
- Fischer, G., Wenner, T., Decaris, B., and Leblond, P. (1998). Chromosomal arm replacement generates a high level of intraspecific polymorphism in the terminal inverted repeats of the linear chromosomal DNA of *Streptomyces ambofaciens*. *Proc. Natl. Acad. Sci. U. S. A.* 95, 14296–14301.
- Frit, P., Barboule, N., Yuan, Y., Gomez, D., and Calsou, P. (2014). Alternative end-joining pathway(s): Bricolage at DNA breaks. *DNA Repair* 17, 81–97. doi:10.1016/j.dnarep.2014.02.007.
- Furuya, E. Y., and Lowy, F. D. (2006). Antimicrobial-resistant bacteria in the community setting. *Nature Rev. Microbiol.* 4, 36–45. doi:10.1038/nrmicro1325.
- Gamulin, V., Cetkovic, H., and Ahel, I. (2004). Identification of a promoter motif regulating the major DNA damage response mechanism of. *FEMS Microbiol. Lett.* 238, 57–63. doi:10.1016/j.femsle.2004.07.017.
- Gong, C., Bongiorno, P., Martins, A., Stephanou, N. C., Zhu, H., Shuman, S., *et al.* (2005). Mechanism of nonhomologous end-joining in mycobacteria: a low-fidelity repair system driven by Ku, ligase D and ligase C. *Nat. Struct. Mol. Biol.* 12, 304–312. doi:10.1038/nsmb915.
- Gong, C., Martins, A., Bongiorno, P., Glickman, M., and Shuman, S. (2004). Biochemical and Genetic Analysis of the Four DNA Ligases of Mycobacteria. *J. Biol. Chem.* 279, 20594–20606. doi:10.1074/jbc.M401841200.
- Gordon, J. L., Byrne, K. P., and Wolfe, K. H. (2011). Mechanisms of chromosome number evolution in yeast. *PLoS Genet.* 7, e1002190.
- Gregory, M. A., Till, R., and Smith, M. C. M. (2003). Integration Site for *Streptomyces* Phage BT1 and Development of Site-Specific Integrating Vectors. *J. Bacteriol.* 185, 5320–5323. doi:10.1128/JB.185.17.5320-5323.2003.
- Groth, A. C., and Calos, M. P. (2004). Phage Integrases: Biology and Applications. *J. Mol. Biol.* 335, 667–678. doi:10.1016/j.jmb.2003.09.082.

- Gupta, R., Barkan, D., Redelman-Sidi, G., Shuman, S., and Glickman, M. S. (2011). Mycobacteria exploit three genetically distinct DNA double-strand break repair pathways. *Mol. Microbiol.* 79, 316–330. doi:10.1111/j.1365-2958.2010.07463.x.
- Gust, B., Challis, G. L., Fowler, K., Kieser, T., and Chater, K. F. (2003). PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proc. Natl. Acad. Sci. U. S. A.* 100, 1541–1546.
- Hargreaves, D., Rafferty, J. B., Sedelnikova S. E., Lloyd R. G., and Rice D. W. (1999). Crystallization of *Escherichia coli* RuvA complexed with a synthetic Holliday junction. *Acta Crystallogr. D. Biol. Crystallogr.* 55, 263-265.
- Holliday R. (1964). A mechanism for gene conversion in fungi. *Genet. Res.* 5, 282-304.
- Huang, T.-W., and Chen, C. W. (2006). A recA Null Mutation May Be Generated in *Streptomyces coelicolor*. *J. Bacteriol.* 188, 6771–6779. doi:10.1128/JB.00951-06.
- Huang, C.-H., Lin, Y.-S., Yang, Y.-L., Huang, S., and Chen, C. W. (1998). The telomeres of *Streptomyces* chromosomes contain conserved palindromic sequences with potential to form complex secondary structures. *Mol. Microbiol.* 28, 905–916.
- Ikeda, H., Ishikawa, J., Hanamoto, A., Shinose, M., Kikuchi, H., Shiba, T., *et al.* (2003). Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. *Nature Biotechnol.* 21, 526–531. doi:10.1038/nbt820.
- Inoue, S., Higashiyama, K., Uchida, T., Hiratsu, K., and Kinashi, H. (2003) Chromosomal circularization in *Streptomyces griseus* by nonhomologous recombination of deletion ends. *Biosci. Biotechnol. Biochem.* 67(5), 1101-1108.
- Iwasaki, H., Takahagi, M., Shiba, T., Nakata, A., and Shinagawa, H. (1991). *Escherichia coli* RuvC protein is an endonuclease that resolves the Holliday structure. *EMBO J.* 10, 4381.
- Jacquier, A., and Dujon, B. (1985) An Intron-Encoded Protein is Active in a Gene Conversion Process That Spreads an Intron into a Mitochondrial Gene. *Cell* 41, 383-394
- Jinks-Robertson, S., and Petes, T. D. (1985). High frequency meiotic gene conversion between repeated genes on nonhomologous chromosomes in yeast. *Proc. Natl. Acad. Sci. U. S. A.* 82, 3350-3354.
- Kameoka, D., Lezhava, A., Zenitani, H., Hiratsu, K., Kawamoto, M., Goshi, K., *et al.* (1999). Analysis of fusion junctions of circularized chromosomes in *Streptomyces griseus*. *J. Bacteriol.* 181, 5711–5717.
- Kaplan, D. L., and O'Donnell, M. (2002). DnaB drives DNA branch migration and dislodges proteins while encircling two DNA strands. *Mol. Cell* 10, 647–657.
- Kim, J.-N., Kim, Y., Jeong, Y., Roe, J.-H., Kim, B.-G., and Cho, B.-K. (2015). Comparative Genomics Reveals the Core and Accessory Genomes of *Streptomyces* Species. *J. Microbiol. Biotechnol.* 25, 1599–1605. doi:10.4014/jmb.1504.04008.

- Kobayashi, H., Simmons, L. A., Yuan, D. S., Broughton, W. J., and Walker, G. C. (2007). Multiple Ku orthologues mediate DNA non-homologous end-joining in the free-living form and during chronic infection of *Sinorhizobium meliloti*. *Mol. Microbiol.* 67, 350–363. doi:10.1111/j.1365-2958.2007.06036.x.
- Kogoma, T. (1997). Stable DNA replication: interplay between DNA replication, homologous recombination, and transcription. *Microbiol. Mol. Biol. Rev.* 61, 212–238.
- Kogoma, T., Cadwell, G. W., Barnard, K. G., and Asai, T. (1996). The DNA replication priming protein, PriA, is required for homologous recombination and double-strand break repair. *J. Bacteriol.* 178, 1258–1264.
- Krokan, H. E., and Bjoras, M. (2013). Base Excision Repair. *Cold Spring Harb. Persp. Biol.* 5, a012583–a012583. doi:10.1101/cshperspect.a012583.
- Laureti, L., Song, L., Huang, S., Corre, C., Leblond, P., Challis, G. L., *et al.* (2011). Identification of a bioactive 51-membered macrolide complex by activation of a silent polyketide synthase in *Streptomyces ambofaciens*. *Proc. Natl. Acad. Sci. U. S. A.* 108, 6258–6263. doi:10.1073/pnas.1019077108.
- Lawrence, J. G., and Ochman, H. (1998). Molecular archaeology of the *Escherichia coli* genome. *Proc. Natl. Acad. Sci. U. S. A.* 95, 9413–9417.
- Leblond, P., Demuyter, P., Moutier, L., Laakel, M., Decaris, B., and Simonet, J.-M. (1989). Hypervariability, a new phenomenon of genetic instability, related to DNA amplification in *Streptomyces ambofaciens*. *J. Bacteriol.* 171, 419–423.
- Leblond, P., Fischer, G., Francou, F.-X., Berger, F., Guérineau, M., and Decaris, B. (1996). The unstable region of *Streptomyces ambofaciens* includes 210 kb terminal inverted repeats flanking the extremities of the linear chromosomal DNA. *Mol. Microbiol.* 19, 261–271.
- Lee, H.-H., Hsu, C.-C., Lin, Y.-L., and Chen, C. W. (2011). Linear plasmids mobilize linear but not circular chromosomes in *Streptomyces*: support for the “end first” model of conjugal transfer. *Microbiol.* 157, 2556–2568. doi:10.1099/mic.0.051441-0.
- Lee, S. E., Moore, J. K., Holmes, A., Umezu, K., Kolodner, R. D., and Haber, J. E. (1998). *Saccharomyces* Ku70, Mre11/Rad50, and RPA proteins regulate adaptation to G2/M arrest after DNA damage. *Cell* 94, 399–409.
- Li, Z., Wen, J., Lin, Y., Wang, S., Xue, P., Zhang, Z., *et al.* (2011). A Sir2-like protein participates in mycobacterial NHEJ. *PloS one* 6, e20045.
- Lieber, M. R. (2008). The Mechanism of Human Nonhomologous DNA End Joining. *J. Biol. Chem.* 283, 1–5. doi:10.1074/jbc.R700039200.
- Lieber, M. R. (2010). The Mechanism of Double-Strand DNA Break Repair by the Nonhomologous DNA End-Joining Pathway. *Annu. Rev. Biochem.* 79, 181–211. doi:10.1146/annurev.biochem.052308.093131.
- Lieber, M. R., Ma, Y., Pannicke, U., and Schwarz, K. (2003). Mechanism and regulation of

- human non-homologous DNA end-joining. *Nat. Rev. Mol. Cell Biol.* 4, 712-720.
- Lin, Z., Kong, H., Nei, M., and Ma, H. (2006). Origins and evolution of the *recA/RAD51* gene family: evidence for ancient gene duplication and endosymbiotic gene transfer. *Proc. Natl. Acad. Sci. U. S. A.* 103, 10328–10333.
- Liu, L., Chen, X., Skogerbø, G., Zhang, P., Chen, R., He, S., et al. (2012). The human microbiome: A hot spot of microbial horizontal gene transfer. *Genom.* 100, 265–270. doi:10.1016/j.ygeno.2012.07.012.
- Lloyd, R. G., and Buckman, C. (1991). Genetic analysis of the *recG* locus of *Escherichia coli* K-12 and of its role in recombination and DNA repair. *J. Bacteriol.* 173, 1004–1011.
- Lloyd, R. G., and Rudolph, C. J. (2016). 25 years on and no end in sight: a perspective on the role of RecG protein. *Curr. Genet.* 62, 827–840. doi:10.1007/s00294-016-0589-z.
- Lloyd, R. G., and Sharples, G. J. (1993a). Dissociation of synthetic Holliday junctions by *E. coli* RecG protein. *EMBO J.* 12, 17.
- Lloyd, R. G., and Sharples, G. J. (1993b). Processing of recombination intermediates by the RecG and RuvAB proteins of *Escherichia coli*. *Nucleic Acids Res.* 21, 1719–1725.
- Magner, D. B., Blankschien, M. D., Lee, J. A., Pennington, J. M., Lupski, J. R., and Rosenberg, S. M. (2007). RecQ Promotes Toxic Recombination in Cells Lacking Recombination Intermediate-Removal Proteins. *Mol. Cell* 26, 273–286. doi:10.1016/j.molcel.2007.03.012.
- Mahdi, A. A., Sharples, G. J., Mandal, T. N., and Lloyd, R. G. (1996). Holliday Junction Resolvases Encoded by Homologous *rusA* Genes in *Escherichia coli* K-12 and Phage 82. *J. Mol. Biol.* 257, 561–573.
- Mascarenhas, J., Sanchez, H., Tadesse, S., Kidane, D., Krisnamurthy, M., Alonso, J. C., et al. (2006). *Bacillus subtilis* SbcC protein plays an important role in DNA inter-strand cross-link repair. *BMC Mol. Biol.* 7, 1.
- Masterson, C., Boehmer, P. E., McDonald, F., Chaudhuri, S., Hickson, I. D., and Emmerson, P. T. (1992). Reconstitution of the activities of the RecBCD holoenzyme of *Escherichia coli* from the purified subunits. *J. Biol. Chem.* 267, 13564–13572.
- Mathiasen, D. P., and Lisby, M. (2014). Cell cycle regulation of homologous recombination in *Saccharomyces cerevisiae*. *FEMS Microbiol. Rev.* 38, 172–184. doi:10.1111/1574-6976.12066.
- Matic, I., Rayssiguier, C., and Radman, M. (1995) -Interspecies Gene Exchange in Bacteria, the Role of SOS and Mismatch Repair Systems in Evolution of Species. *Cell* 80, 507-515
- McGovern, S., Baconnais, S., Roblin, P., Nicolas, P., Drevet, P., Simonson, H., et al. (2016). C-terminal region of bacterial Ku controls DNA bridging, DNA threading and recruitment of DNA ligase D for double strand breaks repair. *Nucleic Acids Res.* 44, 4785–4806. doi:10.1093/nar/gkw149.
- Meddows, T. R., Savory, A. P., and Lloyd, R. G. (2004). RecG helicase promotes DNA double-

- strand break repair: DNA double-strand break repair. *Mol. Microbiol.* 52, 119–132. doi:10.1111/j.1365-2958.2003.03970.x.
- Mehta, A., and Haber, J. E. (2014). Sources of DNA Double-Strand Breaks and Models of Recombinational DNA Repair. *Cold Spring Harb. Perspect. Biol.* 6, a016428–a016428. doi:10.1101/cshperspect.a016428.
- Michel, B., and Ehrlich, S. D. (1984). Recombination efficiency is a quadratic function of the length of homology during plasmid transformation of *Bacillus subtilis* protoplasts and *Escherichia coli* competent cells. *EMBO J.* 3, 2879.
- Michel, B., and Leach, D. (2012). Homologous Recombination—Enzymes and Pathways. *EcoSal Plus* 5. doi:10.1128/ecosalplus.7.2.7.
- Michel, B., Niaudet, B., and Ehrlich, S. D. (1982). Intramolecular recombination during plasmid transformation of *Bacillus subtilis* competent cells. *EMBO J.* 1, 1565.
- Mimitou, E. P., and Symington, L. S. (2010). Ku prevents Exo1 and Sgs1-dependent resection of DNA ends in the absence of a functional MRX complex or Sae2. *EMBO J.* 29, 3358–3369.
- Moeller, R., Stackebrandt, E., Reitz, G., Berger, T., Rettberg, P., Doherty, A. J., et al. (2007). Role of DNA Repair by Nonhomologous-End Joining in *Bacillus subtilis* Spore Resistance to Extreme Dryness, Mono- and Polychromatic UV, and Ionizing Radiation. *J. Bacteriol.* 189, 3306–3311. doi:10.1128/JB.00018-07.
- Moore, J. K., and Haber, J. E. (1996). Cell cycle and genetic requirements of two pathways of nonhomologous end-joining repair of double-strand breaks in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 16, 2164–2173.
- Müller, B., and West, S. C. (1994). Processing of Holliday junctions by the *Escherichia coli* RuvA, RuvB, RuvC and RecG proteins. *Experientia* 50, 216–222.
- Myronovskyi, M., Tokovenko, B., Brötz, E., Rückert, C., Kalinowski, J., and Luzhetskyy, A. (2014). Genome rearrangements of *Streptomyces albus* J1074 lead to the carotenoid gene cluster activation. *Appl. Microbiol. Biotechnol.* 98, 795–806. doi:10.1007/s00253-013-5440-6.
- Nautiyal, A., Rani, P. S., Sharples, G. J., and Muniyappa, K. (2016). *Mycobacterium tuberculosis* RuvX is a Holliday junction resolvase formed by dimerisation of the monomeric YqgF nuclease domain. *Mol. Microbiol.* 100, 656–674. doi:10.1111/mmi.13338.
- de Ory, A., Nagler, K., Carrasco, B., Raguse, M., Zafra, O., Moeller, R., et al. (2016). Identification of a conserved 5'-dRP lyase activity in bacterial DNA repair ligase D and its potential role in base excision repair. *Nucleic Acids Res.* 44, 1833–1844. doi:10.1093/nar/gkw054.
- de Ory, A., Zafra, O., and de Vega, M. (2014). Efficient processing of abasic sites by bacterial nonhomologous end-joining Ku proteins. *Nucleic Acids Res.* 42, 13082–13095. doi:10.1093/nar/gku1029.

- Pâques, F., and Haber, J. E. (1999). Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. *Microbiol. Mol. Biol. Rev.* 63, 349-404.
- Paris, Ü., Mikkel, K., Tavita, K., Saumaa, S., Teras, R., and Kivisaar, M. (2015). NHEJ enzymes LigD and Ku participate in stationary-phase mutagenesis in *Pseudomonas putida*. *DNA Repair* 31, 11-18. doi:10.1016/j.dnarep.2015.04.005.
- Parsons, C. A., Tsaneva, I., Lloyd, R. G., and West, S. C. (1992). Interaction of *Escherichia coli* RuvA and RuvB proteins with synthetic Holliday junctions. *Proc. Natl. Acad. Sci. U. S. A.* 89, 5452-5456.
- Persky, N. S., and Lovett, S. T. (2008). Mechanisms of Recombination: Lessons from *E. coli*. *Crit. Rev. Biochem. Mol. Biol.* 43, 347-370. doi:10.1080/10409230802485358.
- Peter, M., and Lloyd, R. G. (1999). RecG helicase activity at three- and four-strand DNA structures. *Nucleic Acids Res.* 27, 3049-3056.
- Pitcher, R. S., Brissett, N. C., and Doherty, A. J. (2007a). Nonhomologous End-Joining in Bacteria: A Microbial Perspective. *Annu. Rev. Microbiol.* 61, 259-282. doi:10.1146/annurev.micro.61.080706.093354.
- Pitcher, R. S., Green, A. J., Brzostek, A., Korycka-Machala, M., Dziadek, J., and Doherty, A. J. (2007b). NHEJ protects mycobacteria in stationary phase against the harmful effects of desiccation. *DNA Repair* 6, 1271-1276. doi:10.1016/j.dnarep.2007.02.009.
- Pitcher, R. S., Tonkin, L. M., Green, A. J., and Doherty, A. J. (2005). Domain Structure of a NHEJ DNA Repair Ligase from *Mycobacterium tuberculosis*. *J. Mol. Biol.* 351, 531-544. doi:10.1016/j.jmb.2005.06.038.
- te Poele, E. M., Bolhuis, H., and Dijkhuizen, L. (2008). Actinomycete integrative and conjugative elements. *Antonie van Leeuwenhoek* 94, 127-143. doi:10.1007/s10482-008-9255-x.
- Possoz, C., Ribard, C., Gagnat, J., Pernodet, J.-L., and Guérineau, M. (2001). The integrative element pSAM2 from *Streptomyces*: kinetics and mode of conjugal transfer. *Mol. Microbiol.* 42, 159-166.
- Raeseide, C., Gaffe, J., Deatherage, D. E., Tenailon, O., Briska, A. M., Ptashkin, R. N., et al. (2014). Large Chromosomal Rearrangements during a Long-Term Evolution Experiment with *Escherichia coli*. *mBio* 5, e01377-14-e01377-14. doi:10.1128/mBio.01377-14.
- Rafferty, J. B., Ingleston, S. M., Hargreaves, D., Artymiuk, P. J., Sharples, G. J., Lloyd, R. G., et al. (1998). Structural similarities between *Escherichia coli* RuvA protein and other DNA-binding proteins and a mutational analysis of its binding to the Holliday junction. *J. Mol. Biol.* 278, 105-116.
- Rand, L., Hinds, J., Springer, B., Sander, P., Buxton, R. S., and Davis, E. O. (2003). The majority of inducible DNA repair genes in *Mycobacterium tuberculosis* are induced independently of RecA. *Mol. Microbiol.* 50, 1031-1042. doi:10.1046/j.1365-2958.2003.03765.x.

- Ricchetti, M., Dujon, B., and Fairhead, C. (2003). Distance from the Chromosome End Determines the Efficiency of Double Strand Break Repair in Subtelomeres of Haploid Yeast. *J. Mol. Biol.* 328, 847–862. doi:10.1016/S0022-2836(03)00315-2.
- Rocha, E. P. C., Cornet, E., and Michel, B. (2005). Comparative and Evolutionary Analysis of the Bacterial Homologous Recombination Systems. *PLoS Genet.* 1, e15. doi:10.1371/journal.pgen.0010015.
- Rouse, J. (2009). Control of genome stability by Slx protein complexes. *Biochem. Soc. Trans.* 37, 495–510. doi:10.1042/BST0370495.
- Sanchez, H. Kidane D., Reed, P., Curtis F. A., Castillo Cozar, M., Graumann P. L., *et al.* (2005). The RuvAB Branch Migration Translocase and RecU Holliday Junction Resolvase Are Required for Double-Stranded DNA Break Repair in *Bacillus subtilis*. *Genet.* 171, 873–883. doi:10.1534/genetics.105.045906.
- Sanchez, H., Carrasco, B., Cozar, M. C., and Alonso, J. C. (2007). *Bacillus subtilis* RecG branch migration translocase is required for DNA repair and chromosomal segregation. *Mol. Microbiol.* 65, 920–935. doi:10.1111/j.1365-2958.2007.05835.x.
- Sharples, G. J., Ingleston, S. M., and Lloyd, R. G. (1999). Holliday Junction Processing in Bacteria: Insights from the Evolutionary Conservation of RuvABC, RecG, and RuvA. *J. Bacteriol.* 181, 5543–5550
- Shuman, S., and Glickman, M. S. (2007). Bacterial DNA repair by non-homologous end joining. *Nat. Rev. Microbiol.* 5, 852–861. doi:10.1038/nrmicro1768.
- Siegl, T., Petzke, L., Welle, E., and Luzhetskyy, A. (2010). I-SceI endonuclease: a new tool for DNA repair studies and genetic manipulations in streptomycetes. *Appl. Microbiol. Biotechnol.* 87, 1525–1532. doi:10.1007/s00253-010-2643-y.
- Simsek, D., Brunet, E., Wong, S. Y.-W., Katyal, S., Gao, Y., McKinnon, P. J., *et al.* (2011). DNA Ligase III Promotes Alternative Nonhomologous End-Joining during Chromosomal Translocation Formation. *PLoS Genet.* 7, e1002080. doi:10.1371/journal.pgen.1002080.
- Sinha, K. M., Stephanou, N. C., Gao, F., Glickman, M. S., and Shuman, S. (2007). Mycobacterial UvrD1 Is a Ku-dependent DNA Helicase That Plays a Role in Multiple DNA Repair Events, Including Double-strand Break Repair. *J. Biol. Chem.* 282, 15114–15125. doi:10.1074/jbc.M701167200.
- Sinha, K. M., Unciuleac, M.-C., Glickman, M. S., and Shuman, S. (2009). AdnAB: a new DSB-resecting motor-nuclease from mycobacteria. *Genes Dev.* 23, 1423–1437. doi:10.1101/gad.1805709.
- Spies, M., Amitani, I., Baskin, R. J., and Kowalczykowski, S. C. (2007). RecBCD Enzyme Switches Lead Motor Subunits in Response to χ Recognition. *Cell* 131, 694–705. doi:10.1016/j.cell.2007.09.023.
- Spies 2004 - Homologous recombination by the RecBCD and RecF pathways. (2004).

- Stark, J. M., Pierce, A. J., Oh, J., Pastink, A., and Jasin, M. (2004). Genetic Steps of Mammalian Homologous Repair with Distinct Mutagenic Consequences. *Mol. Cell. Biol.* 24, 9305–9316. doi:10.1128/MCB.24.21.9305-9316.2004.
- Stephanou, N. C., Gao, F., Bongiorno, P., Ehrt, S., Schnappinger, D., Shuman, S., *et al.* (2007). Mycobacterial Nonhomologous End Joining Mediates Mutagenic Repair of Chromosomal Double-Strand DNA Breaks. *J. Bacteriol.* 189, 5237–5246. doi:10.1128/JB.00332-07.
- Stuttard, C. (1982). Temperate phages of *Streptomyces venezuelae*: lysogeny and host specificity shown by phages SV1 and SV2. *Microbiol.* 128, 115–121.
- Svendsen, J. M., and Harper, J. W. (2010). GEN1/Yen1 and the SLX4 complex: solutions to the problem of Holliday junction resolution. *Genes Dev.* 24, 521–536. doi:10.1101/gad.1903510.
- Szostak, J. W., Orr-Weaver, T. L., Rothstein, R. J., and Stahl F. W. (1983) The Double Strand Break Repair Model for Recombination. (1983) *Cell* 33, 25-35.
- Tamas, I., Klasson, L., Canbäck, B., Näslund, A. K., Eriksson, A. S., Wernegreen, J. J. *et al.* (2002). 50 Million Years of Genomic Stasis in Endosymbiotic Bacteria. *Science* 296, 2376–2379. doi:10.1126/science.1071278.
- Thibessard, A., Bertrand, C., Hiblot, J., Piotrowski, E., and Leblond, P. (2015). Construction of pDYN6902, a new *Streptomyces* integrative expression vector designed for cloning sequences interfering with *Escherichia coli* viability. *Plasmid* 82, 43–49. doi:10.1016/j.plasmid.2015.10.003.
- Thibessard, A., Haas, D., Gerbaud, C., Aigle, B., Lautru, S., Pernodet, J.-L., *et al.* (2015). Complete genome sequence of *Streptomyces ambofaciens* ATCC 23877, the spiramycin producer. *J. Biotechnol.* 214, 117–118. doi:10.1016/j.jbiotec.2015.09.020.
- Thibessard, A., and Leblond, P. (2016). Complete Genome Sequence of *Streptomyces ambofaciens* DSM 40697, a Paradigm for Genome Plasticity Studies. *Genome Announc.* 4, e00470-16. doi:10.1128/genomeA.00470-16.
- Thoma, L., and Muth, G. (2012). Conjugative DNA transfer in *Streptomyces* by TraB: is one protein enough? *FEMS Microbiol. Lett.* 337, 81–88. doi:10.1111/1574-6968.12031.
- Thoma, L., and Muth, G. (2015). The conjugative DNA-transfer apparatus of *Streptomyces*. *Int. J. Med. Microbiol.* 305, 224–229. doi:10.1016/j.ijmm.2014.12.020.
- Tong, Y., Charusanti, P., Zhang, L., Weber, T., and Lee, S. Y. (2015). CRISPR-Cas9 Based Engineering of Actinomycetal Genomes. *ACS Synth. Biol.* 4, 1020–1029. doi:10.1021/acssynbio.5b00038.
- Tsai, H.-H., Huang, C.-H., Tessmer, I., Erie, D. A., and Chen, C. W. (2011). Linear *Streptomyces* plasmids form superhelical circles through interactions between their terminal proteins. *Nucleic Acids Res.* 39, 2165–2174. doi:10.1093/nar/gkq1204.
- Uchida, T., Miyawaki, M., and Kinashi, H. (2003). Chromosomal Arm Replacement in *Streptomyces griseus*. *J. Bacteriol.* 185, 1120–1124. doi:10.1128/JB.185.3.1120-1124.2003.

- Unciuleac, M.-C., and Shuman, S. (2010). Characterization of the Mycobacterial AdnAB DNA Motor Provides Insights into the Evolution of Bacterial Motor-Nuclease Machines. *J. Biol. Chem.* 285, 2632–2641. doi:10.1074/jbc.M109.076133.
- Veaute, X., Delmas, S., Selva, M., Jeusset, J., Le Cam, E., Matic, I., *et al.* (2005). UvrD helicase, unlike Rep helicase, dismantles RecA nucleoprotein filaments in *Escherichia coli*. *EMBO J.* 24, 180–189.
- de Vega, M. (2013). The minimal *Bacillus subtilis* nonhomologous end joining repair machinery. *PLoS one* 8, e64232.
- Vierling, S., Weber, T., Wohlleben, W., and Muth, G. (2001). Evidence that an Additional Mutation Is Required To Tolerate Insertional Inactivation of the *Streptomyces lividans recA* Gene. *J. Bacteriol.* 183, 4374–4381. doi:10.1128/JB.183.14.4374-4381.2001.
- Vincent, S. D., Mahdi, A. A., and Lloyd, R. G. (1996). The RecG Branch Migration Protein of *Escherichia coli* Dissociates R-loops. *J. Mol. Biol.* 264, 713–721.
- Vogelmann, J., Ammelburg, M., Finger, C., Guezguez, J., Linke, D., Flötenmeyer, M., *et al.* (2011). Conjugal plasmid transfer in *Streptomyces* resembles bacterial chromosome segregation by FtsK/SpoIIIE. *EMBO J.* 30, 2246–2254. doi:10.1038/emboj.2011.121.
- Volff, J.-N., and Altenbuchner, J. (2000). A new beginning with new ends: linearisation of circular chromosomes during bacterial evolution. *FEMS Microbiol. Lett.* 186, 143–150.
- Wang, S. T., Setlow, B., Conlon, E. M., Lyon, J. L., Imamura, D., Sato, T., *et al.* (2006). The Forespore Line of Gene Expression in *Bacillus subtilis*. *J. Mol. Biol.* 358, 16–37. doi:10.1016/j.jmb.2006.01.059.
- Wardrope, L., Okely, E., and Leach, D. (2009). Resolution of Joint Molecules by RuvABC and RecG Following Cleavage of the *Escherichia coli* Chromosome by EcoKI. *PLoS ONE* 4, e6542. doi:10.1371/journal.pone.0006542.
- Watt, V. M., Ingles, C. J., Urdea, M. S., and Rutter, W. J. (1985). Homology requirements for recombination in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* 82, 4768–4772.
- Weller, G. R. (2002). Identification of a DNA Nonhomologous End-Joining Complex in Bacteria. *Science* 297, 1686–1689. doi:10.1126/science.1074584.
- Wenner, T., Roth, V., Fischer, G., Fourrier, C., Aigle, B., Decaris, B., *et al.* (2003). End-to-end fusion of linear deleted chromosomes initiates a cycle of genome instability in *Streptomyces ambofaciens*. *Mol. Microbiol.* 50, 411–425. doi:10.1046/j.1365-2958.2003.03698.x.
- Whitby, M. C., and Lloyd, R. G. (1998). Targeting Holliday junctions by the RecG branch migration protein of *Escherichia coli*. *J. Biol. Chem.* 273, 19729–19739.
- Whitby, M. C., Ryder, L., and Lloyd, R. G. (1993). Reverse branch migration of Holliday junctions by RecG protein: a new mechanism for resolution of intermediates in recombination and DNA repair. *Cell* 75, 341–350.

- Widenbrant, E. M., Tsai, H.-H., Chen, C. W., and Kao, C. M. (2007). *Streptomyces coelicolor* Undergoes Spontaneous Chromosomal End Replacement. *J. Bacteriol.* 189, 9117–9121. doi:10.1128/JB.01049-07.
- Widenbrant, E. M., Tsai, H.-H., Chen, C. W., and Kao, C. M. (2008). Spontaneous Amplification of the Actinorhodin Gene Cluster in *Streptomyces coelicolor* Involving Native Insertion Sequence IS466. *J. Bacteriol.* 190, 4754–4758. doi:10.1128/JB.00131-08.
- Wozniak, R. A. F., and Waldor, M. K. (2010). Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow. *Nature Rev. Microbiol.* 8, 552–563. doi:10.1038/nrmicro2382.
- Wu, D., Topper, L. M., and Wilson, T. E. (2008). Recruitment and Dissociation of Nonhomologous End Joining Proteins at a DNA Double-Strand Break in *Saccharomyces cerevisiae*. *Genet.* 178, 1237–1249. doi:10.1534/genetics.107.083535.
- Yanai, K., Murakami, T., and Bibb, M. (2006). Amplification of the entire kanamycin biosynthetic gene cluster during empirical strain improvement of *Streptomyces kanamyceticus*. *Proc. Natl. Acad. Sci. U. S. A.* 103, 9661–9666.
- Yeeles, J. T. P., and Dillingham, M. S. (2007). A Dual-nuclease Mechanism for DNA Break Processing by AddAB-type Helicase-nucleases. *J. Mol. Biol.* 371, 66–78. doi:10.1016/j.jmb.2007.05.053.
- Young, M., and Cullum, J. (1987). A plausible mechanism for large-scale chromosomal DNA amplification in streptomycetes. *FEBS Lett.* 212, 10–14.
- Zegeye, E. D., Balasingham, S. V., Laerdahl, J. K., Homberset, H., and Tonjum, T. (2012). *Mycobacterium tuberculosis* RecG binds and unwinds model DNA substrates with a preference for Holliday junctions. *Microbiol.* 158, 1982–1993. doi:10.1099/mic.0.058693-0.
- Zhang, L., Nguyen, H. C., Chipot, L., Piotrowski, E., Bertrand, C., Thibessard, A., et al. (2014). The *adnAB* Locus, Encoding a Putative Helicase-Nuclease Activity, Is Essential in *Streptomyces*. *J. Bacteriol.* 196, 2701–2708. doi:10.1128/JB.01513-14.
- Zhang, X., Chen, W., Zhang, Y., Jiang, L., Chen, Z., Wen, Y., et al. (2012). Deletion of *ku* homologs increases gene targeting frequency in *Streptomyces avermitilis*. *J. Ind. Microbiol. Biotechnol.* 39, 917–925. doi:10.1007/s10295-012-1097-x.
- Zhang, Y., Hefferin, M. L., Chen, L., Shim, E. Y., Tseng, H.-M., Kwon, Y., et al. (2007). Role of Dnl4–Lif1 in nonhomologous end-joining repair complex assembly and suppression of homologous recombination. *Nat. Struct. Mol. Biol.* 14, 639–646. doi:10.1038/nsmb1261.
- Zhong, W., Feng, H., Santiago, F. E., and Kipreos, E. T. (2003). CUL-4 ubiquitin ligase maintains genome stability by restraining DNA-replication licensing. *Nature* 423, 885–889. doi:10.1038/nature01747.
- Zhou, Y., Chen, T., Zhou, L., Fleming, J., Deng, J., Wang, X., et al. (2015). Discovery and characterization of Ku acetylation in *Mycobacterium smegmatis*. *FEMS Microbiol. Lett.* 362,

fnu051-fnu051. doi:10.1093/femsle/fnu051.

Zhu, H., Bhattarai, H., Yan, H.-G., Shuman, S., and Glickman, M. S. (2012). Characterization of *Mycobacterium smegmatis* PolD2 and PolD1 as RNA/DNA Polymerases Homologous to the POL Domain of Bacterial DNA Ligase D. *Biochem.* 51, 10147–10158. doi:10.1021/bi301202e.

Zhu, H., Nandakumar, J., Aniukwu, J., Wang, L. K., Glickman, M. S., Lima, C. D., et al. (2006). Atomic structure and nonhomologous end-joining function of the polymerase component of bacterial DNA ligase D. *Proc. Natl. Acad. Sci. U. S. A.* 103, 1711–1716.

Zhu, H., and Shuman, S. (2006). Substrate Specificity and Structure-Function Analysis of the 3'-Phosphoesterase Component of the Bacterial NHEJ Protein, DNA Ligase D. *J. Biol. Chem.* 281, 13873–13881. doi:10.1074/jbc.M600055200.

Zhu, H., and Shuman, S. (2007). Characterization of *Agrobacterium tumefaciens* DNA ligases C and D. *Nucleic Acids Res.* 35, 3631–3645. doi:10.1093/nar/gkm145.

Zhu, H., and Shuman, S. (2010). Gap Filling Activities of *Pseudomonas* DNA Ligase D (LigD) Polymerase and Functional Interactions of LigD with the DNA End-binding Ku Protein. *J. Biol. Chem.* 285, 4815–4825. doi:10.1074/jbc.M109.073874.

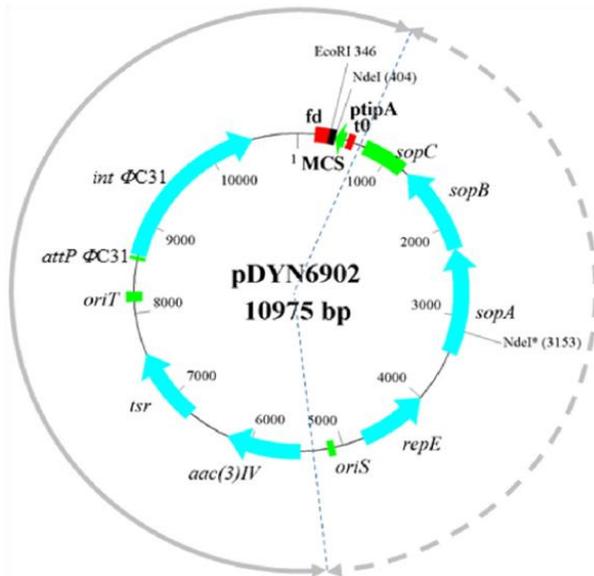
Zhu, H., Wang, L. K., and Shuman, S. (2005). Essential Constituents of the 3'-Phosphoesterase Domain of Bacterial DNA Ligase D, a Nonhomologous End-joining Enzyme. *J. Biol. Chem.* 280, 33707–33715. doi:10.1074/jbc.M506838200.

Zickler, D., and Kleckner, N. (2015). Recombination, Pairing, and Synapsis of Homologs during Meiosis. *Cold Spring Harb. Perspect. Biol.* 7, a016626. doi:10.1101/cshperspect.a016626.

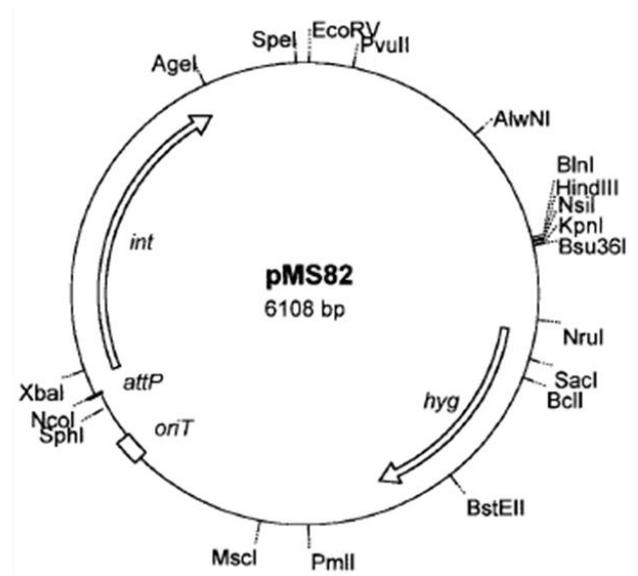
ANNEXES

Annexe 2 : Cartes de restriction du plasmide pDYN6902 (A ; Thibessard *et al.* 2015) et du plasmide pMS82 IKI (B ; Gregory *et al.* 2003)

A)



B)



Résumé :

Rayons ionisants, dessiccation, ou encore métabolites secondaires exogènes sont autant de facteurs qui peuvent engendrer des dommages à l'ADN chez les bactéries du sol, notamment en provoquant la formation de cassures double-brin (DSB), préjudice majeur pour une cellule. Chez les procaryotes, l'évolution a sélectionné deux principaux mécanismes de réparation des DSB, à savoir la recombinaison homologe (RH) et le *non-homologous end joining* (NHEJ). La RH est un mécanisme quasi-ubiquiste dans le monde bactérien qui repose sur l'utilisation d'une copie intacte de la molécule endommagée comme matrice pour la réparation de la DSB. Contrairement à la RH, le NHEJ n'est présent que chez 20 à 25% des bactéries et est considéré comme un mécanisme mutagène puisque la réparation de la DSB se fait sans matrice homologe et peut entraîner l'ajout ou la délétion de nucléotides au site de cassure. Chez la bactérie modèle *Mycobacterium*, seuls deux acteurs sont nécessaires pour la réparation par NHEJ. Ainsi, un dimère de protéine Ku se fixe sur la cassure puis recrute la protéine multifonctionnelle LigD, qui catalyse le traitement puis la ligation des extrémités grâce à ses domaines polymérase, nucléase et ligase.

Les mécanismes de réparation des DSB chez les *Streptomyces* étaient peu connus à l'initiation de ce travail. Cette bactérie présente des caractéristiques génomiques remarquables avec notamment un chromosome linéaire de grande taille (6 à 12 Mb). En ce qui concerne la RH, nous avons focalisé nos recherches sur les étapes tardives (post-synaptiques) et étudié le rôle du complexe RuvABC et de RecG impliqués chez *Escherichia coli* dans la migration de la croix de Holliday et de sa résolution. La construction de mutants simples et multiples a montré que bien que les gènes codant ces protéines soient très conservés chez les *Streptomyces*, leur déficience ne se traduit chez *Streptomyces ambofaciens* que par une faible baisse de la recombinaison suite à un événement de conjugaison. Aucune baisse de l'efficacité de recombinaison intrachromosomique n'a en revanche été observée. Ces résultats suggèrent que des acteurs alternatifs majeurs sont encore à découvrir chez les *Streptomyces*. Le décryptage du mécanisme de NHEJ chez *S. ambofaciens* constitue une première dans ce genre bactérien. Une étude génomique exhaustive a permis de révéler la très grande diversité du nombre d'acteurs potentiels de ce mécanisme (Ku, LigDom, PolDom, NucDom) et de l'organisation des gènes qui les codent. L'analyse fonctionnelle a révélé que l'ensemble des acteurs étaient impliqués dans la réponse à l'exposition à un faisceau d'électrons accélérés, connus pour induire, entre autre, la formation de DSB. La génération de DSB, par coupure endonucléasique I-SceI, a par ailleurs permis de mettre en évidence au niveau moléculaire des réparations de type NHEJ (délétions ou insertions de quelques nucléotides, intégration de fragments d'ADN). Les cassures dans les régions terminales du chromosome sont accompagnées de grandes délétions (jusqu'à 2,1 Mb) et de réarrangements de grande ampleur incluant circularisations du chromosome et amplifications d'ADN. Les conséquences de la réparation de DSB chez *S. ambofaciens* sont en tous points similaires aux réarrangements observés spontanément ou par comparaison des génomes des espèces types. Ainsi, il est possible de lier la plasticité du génome à la réparation de DSB. En outre, l'intégration de matériel génétique exogène serait favorisée au cours de la réparation NHEJ ce qui donnerait à ce système de réparation une place importante dans le processus de transfert horizontal, mécanisme d'évolution majeur chez les bactéries.

Abstract:

Ionizing radiation, desiccation or exogenous secondary metabolites are all factors that can cause DNA damage in soil bacteria, especially by triggering double strand breaks (DSB), the most detrimental harm for the cell. In prokaryotes, evolution selected two main DSB repair pathways, namely homologous recombination (HR) and non-homologous end joining (NHEJ). HR is almost ubiquitous in bacteria and relies on an intact copy of the damaged DNA molecule as a template for DSB repair. In contrast to HR, NHEJ is only present in 20 to 25% of bacteria and is considered as a mutagenic pathway since DSB repair is performed without the need of any template and can lead to nucleotide addition or deletion at DSB site. In the bacterial model *Mycobacterium*, two partners are sufficient for a functional NHEJ pathway. Thus, Ku protein dimer recognizes and binds the DSB and then recruits the multifunctional LigD protein for extremities treatment and ligation thanks to its polymerase, nuclease and ligase domains.

At the beginning of this work, few informations on DSB repair in *Streptomyces* were available. This bacteria exhibits remarkable genomic features including a large linear chromosome (6 to 12 Mb). Regarding HR, we focused on the late stage (post-synaptic step) in studying the role of RuvABC complex and RecG, involved in branch migration and Holliday junction resolution in *E. coli*. Construction of single and multiple mutants showed that although the genes encoding these proteins are highly conserved in *Streptomyces*, their deficiency in *Streptomyces ambofaciens* only results in a mild decrease of recombination after conjugation events. Besides, no decrease of intrachromosomal recombination efficiency could be observed. These results suggest that major alternative factors are still to be discovered in *Streptomyces*. This work was also the first occasion to decipher a NHEJ pathway in *Streptomyces*. An exhaustive genomic study revealed a great diversity in the number of factors potentially implicated in this pathway (Ku, LigDom, PolDom, NucDom) and in the organization of their encoding genes. Functional analyses revealed that all the factors, whatever they are conserved or not between species, were involved in the response to electron beam exposure, known to induce, amongst other things, DSB formation. Generation of DSB by I-SceI endonuclease cleavage was also used to evidence at a molecular level NHEJ type DSB repair (deletions or insertions of several nucleotides, integration of DNA fragments). Targeted breaks in the terminal regions of the chromosome were accompanied by large deletions (up to 2.1 Mb) and major rearrangements including chromosome circularizations and DNA amplifications. Consequences of DSB repair in *S. ambofaciens* are in all points similar to chromosome rearrangements observed spontaneously or by comparing genomes of different species. Thus, it is possible to link the genome plasticity to DSB repair. In addition, the integration of exogenous genetic material would be favoured during NHEJ repair which would give this repair system a major role in the horizontal transfer process, known to be a main evolution mechanism in bacteria.