

Étudier et combattre les bactéries pathogènes

Les outils CRISPR

David Bikard est directeur du laboratoire de Biologie de Synthèse à l'Institut Pasteur. Ingénieur de l'Agro-ParisTech, il a obtenu un doctorat à l'Université Paris-Diderot pour ses travaux réalisés à l'Institut Pasteur sur la recombinaison chez les bactéries. En post-doctorat à la Rockefeller University, il a initié ses recherches sur CRISPR.

1 Qu'est-ce que les CRISPR ?

1.1. Une histoire fascinante et peu connue

L'histoire remonte en 1987 avec un groupe japonais qui a publié la première observation d'un locus¹ « CRISPR » dans un génome de la bactérie *Escherichia coli*² où l'on voit ces séquences répétées qui sont régulièrement espacées par des séquences relativement variées. À l'époque on n'avait aucune idée de ce que

ces séquences pouvaient bien être (**Figure 1**).

Le sens de cette occurrence est resté un point obscur pendant relativement longtemps. Après cette découverte de 1987, des bio-informaticiens ont repéré ces séquences répétées en de nombreux endroits différents (**Figure 2**) à mesure qu'on séquençait les génomes microbiens. Ces découvertes ont été réalisées indépendamment à plusieurs reprises.

À chaque fois ces chercheurs observaient la présence de répétitions (en blanc sur la **Figure 3**). Les espaceurs³ correspondent à des régions

1. Locus : localisation précise d'un gène sur un chromosome.

2. *Escherichia coli* : bactérie intestinale (Gram négatif) des mammifères, très commune chez l'être humain, composant environ 80 % de notre flore intestinale aérobie.

3. Espaceur : région non-codante d'ADN entre les gènes.

Nucleotide Sequence of the *iap* Gene, Responsible for Alkaline Phosphatase Isozyme Conversion in *Escherichia coli*, and Identification of the Gene Product

YOSHIZUMI ISHINO, HIDEO SHINAGAWA, KOZO MAKINO, MITSUKO AMEMURA, AND ATSUO NAKATA*
 Department of Experimental Chemotherapy, The Research Institute for Microbial Diseases, Osaka University, 3-1 Yamadaoka, Suita, Osaka 565, Japan

Figure 1

Publication, en 1987, de la première observation d'un locus CRISPR dans un génome de la bactérie *Escherichia coli*.

GGAGTTCAC	CGCAGAGCGC	GGGGAACTCC	-iap-TGAT	GGGTTGAAA	ATGGGAGCTG	(1390
CGGTTTATCC	CCGCTGATGC	GGGGAACACC	AAGTGATATC	CATCATCGCA	TCCAGTGGCC	C (1451
CGGTTTATCC	CTGCTGGCGC	GGGGAACTCT	AGCGTCAGGC	GTGAAATCTC	ACCGTCGTTG	C (1512
CGGTTTATCC	CCGCTAACGC	GGGGAACTCG	CGGTTCAGGC	GTTGCAAACC	TGGCTACCGG	G (1573
CGGTTTATCC	CCGCTGGCGC	GGGGAACTCG	TAGTCCATCA	TTCACCTAT	GTCTGAACTC	C (1634
CGGTTTATCC	CCGCTGGCGC	GGGGAACTCC	CGGGGATAA	TGTTACGGT	CATGCGCCCC	C (1695
CGGTTTATCC	CCGCTGGCGC	GGGGAACTCT	GGGCGGCTTG	CCTTGCAGCC	AGCTCCAGCA	G (1756
CGGTTTATCC	CCGCTGGCGC	GGGGAACTCA	AGCTGGCTGG	CAATCTCTT	CGGGGTGAGT	C (1817
CGGTTTATCC	CCGCTGGCGC	GGGGAACTCT	AGTTTCCGTA	TCCTCGGATT	TATAAAGCTG	A (1878
CGGTTTATCC	CCGCTGGCGC	GGGGAACTCG	CAGGCGGCGA	CCGGCAGGT	ATGCGCGATT	CG (1940
CGGTTTATCC	CCGCTGGCGC	GGGGAACTCG	CGACCGCTCA	GAATTCAG	ACCCGATCCA	AA (2002
CGGTTTATCC	CCGCTGGCGC	GGGGAACTCT	CAACATTATC	AATTACAACC	GACAGGGAGC	C (2063
CGGTTTATCC	CCGCTGGCGC	GGGGAACTCA	GCGTGTTCCG	CATCACCTTT	GGCTTCGGCT	G (2124
CGGTTTATCC	CCGCTGGCGC	GGGGAACTCT	GCGTGAGCGT	ATCGCCGCGC	GTCTGCGAAA	G (2185
CGGTTTATCC	CCGCTGGCGC	GGGGAACTCT	CTAAAAGTAT	ACATTTGTTT	TTAAAGCATT	(2255

Figure 2

Séquences répétées et séquences variables dans le génome de *E. coli*.



Figure 3

Régions variables sur des gènes (espaceurs, en couleurs) qui espacent des répétitions (en blanc, qui sont en réalité les CRISPR).

variables non identifiées à l'époque mais associés à des gènes, qu'on a appelés « gènes Cas » pour « *CRISPR associated* ». De nombreux acronymes ont eu cours jusqu'en 2002 (DVR, TREP, LTRR), jusqu'à ce qu'on arrive à CRISPR, « *Clustered*

Regularly Interspaced Short Palindromic Repeats ». Ce nom est resté, et tout le monde l'utilise dans le domaine.

Il a fallu attendre 2005 pour qu'on établisse les premières hypothèses correctes sur la fonction de ces locus CRISPR. En 2005, trois laboratoires ont

indépendamment identifié que ces séquences espaceurs rapelaient des séquences d'ADN étrangers aux bactéries, et en particulier des séquences de bactériophages⁴ ou des séquences de plasmides⁵. Trois publications ont émis l'hypothèse qu'il s'agirait d'un système de défense permettant aux bactéries de se défendre contre ces éléments étrangers. Il a fallu attendre encore deux ans, en 2007, et un manuscrit du journal *Science* (Figure 4), pour apporter la première preuve expérimentale que ces systèmes CRISPR sont bien un système immunitaire utilisé par les bactéries pour se défendre contre les bactériophages.

L'histoire de cet article est d'ailleurs intéressante puisqu'il ne provient pas d'un

groupe académique mais d'une industrie, *Danisco*, qui produit des ferments lactiques pour la fabrication de fromages et yaourts. C'est une industrie où sont développées des cultures de milliers de litres de lait, et lorsque ces cultures de bactéries sont contaminées par des bactériophages, qui peuvent tuer toutes les bactéries, la production peut être complètement ruinée. C'est donc une industrie où sont en permanence sélectionnés des variants⁶ résistants aux bactériophages, selon le procédé du « challenge⁷ », pratiqué par exemple par l'équipe de Rodolphe Barrangou et Philippe Horvath (Figure 4).

En réalisant des « challenges » avec des bactériophages, ces chercheurs se sont aperçus qu'ils pouvaient sélectionner des

4. Bactériophage : un bactériophage, ou phage, est un virus n'infectant que des bactéries. Ce sont des outils fondamentaux en biologie de synthèse. Les bactériophages servent entre autres de vecteurs de clonage et de transfert de gènes.

5. Plasmide : désigne une molécule d'ADN distincte de l'ADN chromosomique, capable de réplication autonome et non essentielle à la survie de la cellule.

6. Variant : une mutation génétique ne s'exprimant que dans certaines conditions particulières. Cette mutation peut avoir un effet fonctionnel ou non en fonction du contexte.

7. Challenge : challenge-test est l'étude de l'évolution dans un aliment de populations de micro-organismes inoculés de manière volontaire.

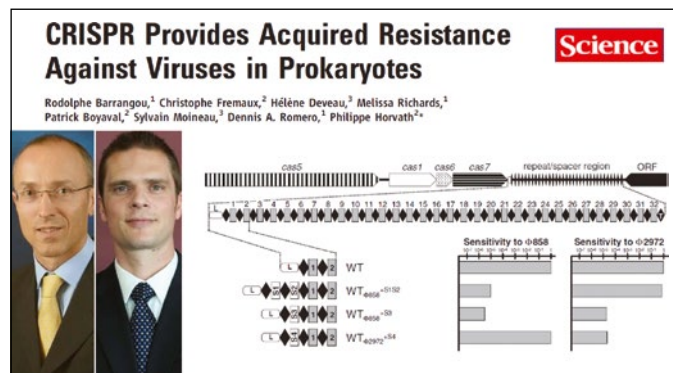


Figure 4

Capture de séquences d'ADN de bactériophages (Φ 858 et Φ 2972) dans les espaceurs de l'ADN de *Streptococcus thermophilus* (WT) par Rodolphe Barrangou et Philippe Horvath.

streptocoques (*Streptococcus thermophilus*) résistants et que lorsqu'ils devenaient résistants, un morceau d'ADN du bactériophage était capturé et intégré sous forme d'un nouvel espaceur au sein du locus CRISPR.

1.2. Comment fonctionnent les systèmes CRISPR dans la nature ?

Pour résumer le fonctionnement du système, on peut distinguer deux phases : une phase d'immunisation et une phase d'immunité.

Dans la **phase d'immunisation**, lorsque le bactériophage attaque la bactérie, il injecte son ADN dans le milieu cellulaire. Par l'intermédiaire de protéines Cas, le système

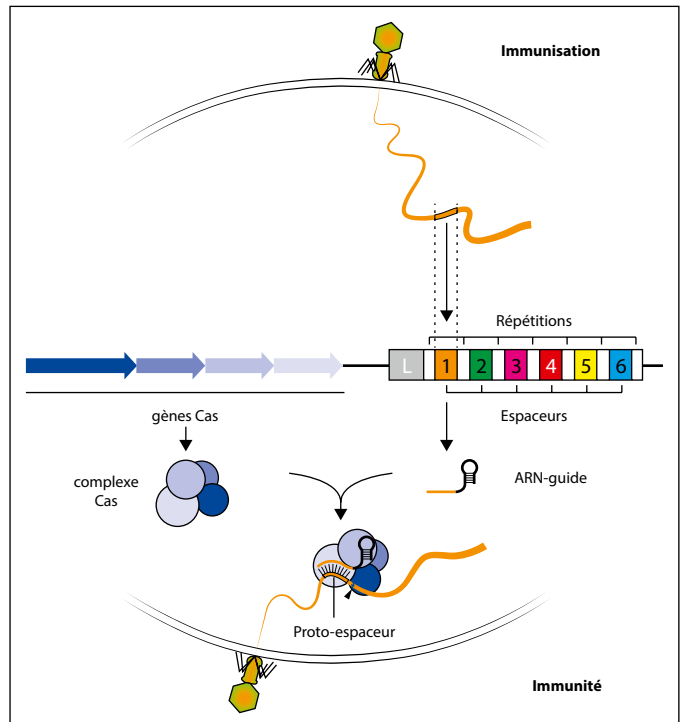
CRISPR capture un morceau de cet ADN du bactériophage et l'intègre sous forme d'un nouvel espaceur entre deux répétitions dans le locus CRISPR. Il construit ainsi une mémoire des infections passées auxquelles il a été soumis.

Cette mémoire est utilisée lors d'une **phase d'immunité** pour combattre les bactériophages homologues à cette séquence ; le locus CRISPR est transcrit en petits ARN-guides⁸ pour guider un complexe de protéines Cas qui vont

8. ARN-guide : ARN qui s'associe à des enzymes protéiques et servent à en guider l'action sur des ARN ou des ADN de séquence complémentaire.

Figure 5

Immunitisation par l'intégration de gènes de bactériophages et immunité par l'expression de ces gènes CRISPR en ARN-guides qui s'associent avec un complexe Cas pour détruire les séquences homologues de pathogènes.



aller détruire les séquences homologues (Figure 5).

Cette découverte est extraordinaire sous plusieurs aspects. Avant cette découverte on ne pensait pas que des êtres unicellulaires pouvaient avoir une immunité adaptative. Par ailleurs, on a réalisé que ces nucléases produits à partir des petits ARN-guides peuvent être à l'origine de nombreuses applications technologiques.

Il existe une énorme diversité de systèmes CRISPR (Figure 6). L'essentiel des applications technologiques qui ont eu lieu jusqu'à présent ont été réalisées avec des systèmes Classe 2 type II, et plus particulièrement type II.A. La Figure 6 représente une classification des différents types de systèmes CRISPR qui travaillent de manière complètement différente, construisant des locus de répétitions et des espaceurs différents puis utilisant ces ARN-guides pour réaliser des choses différentes.

1.3. Le système II.A : applications technologiques

Regardons plus particulièrement le système de type II.A., avec sa protéine Cas9 ; c'est celui qui est utilisé pour toutes les applications technologiques.

On a quatre gènes de protéines Cas : Cas9, Cas1, Cas2 et CSN2. Le locus CRISPR synthétise un ARN précurseur⁹ ainsi qu'un autre petit

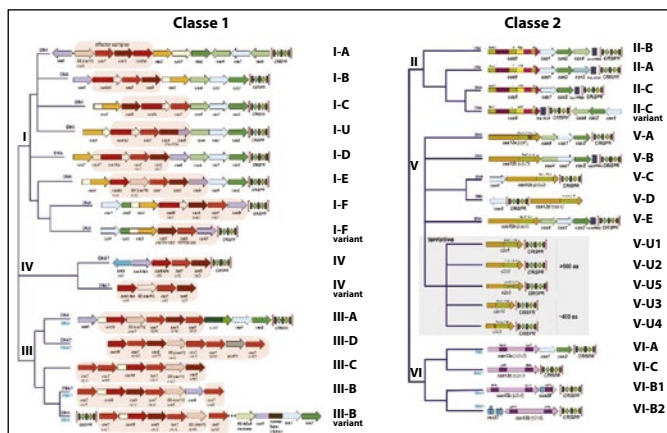


Figure 6

Diversité des systèmes CRISPR.

ARN, qu'on appelle ARN Traceur (pour « *TransActing CRISPR ARN* »). Ce dernier est essentiel au fonctionnement de ce système ; il est homologue aux répétitions de l'ARN précurseur et va s'hybrider pour former l'« ARN duplex », qui sera reconnu d'une part par Cas9 et de l'autre par une ARNase¹⁰. Ce processus conduit à la synthèse des petits ARN-guides (Figure 7).

On obtient, à la fin de ces réactions, un complexe nucléoprotéique comprenant à la fois Cas9 et ces deux ARN. Ce complexe effectue la surveillance dans la cellule ; il est en permanence en train de scanner l'ADN présent dans la bactérie. Lorsqu'il identifie une séquence homologue au guide, il peut détruire cette séquence.

La description de la biochimie du système CRISPR-CAS9 a

9. ARN précurseur : ARN représentant le produit de transcription primaire d'un gène.

10. ARNase : une nucléase qui catalyse la dégradation de l'ARN en éléments plus petits.

été réalisée essentiellement autour de 2012-2013, en particulier par Emmanuelle Charpentier et Jennifer Doudna (Figure 8 : en train de recevoir un prix en compagnie de Cameron Diaz et le CEO de Twitter). Très vite après, on en a eu les premières

applications technologiques. La première application consistait à *modifier les génomes*.

2 Premières utilisations des CRISPR

2.1. Premier détournement de l'immunité adaptative par modification du génome de bactéries : screening et recodage

Très rapidement après la description des systèmes CRISPR-Cas9 chez la bactérie *Streptococcus pyogenes* puis *in vitro*, ce système a été utilisé pour modifier les génomes de bactéries mais également de cellules humaines et de toute une panoplie d'organismes modèles (Figure 9).

À titre d'exemple, la preuve de concept chez la bactérie *E. coli* a été réalisée ainsi : un plasmide qui exprime la protéine Cas9 a d'abord été introduit chez la bactérie, puis un autre plasmide portant un ARN-guide ciblant une position à modifier a été introduit par électroporation¹¹ en même temps qu'un petit morceau d'ADN pouvant s'introduire par recombinaison à l'endroit ciblé (Figure 10). Cette technique permet des modifications extrêmement précises sans laisser de cicatrices et sans marqueurs de sélection sur la cible.

11. Électroporation : technique microbiologique consistant à appliquer un champ électrique sur les membranes cellulaires qui sont ainsi déstabilisées, ce qui augmente la perméabilité membranaire.

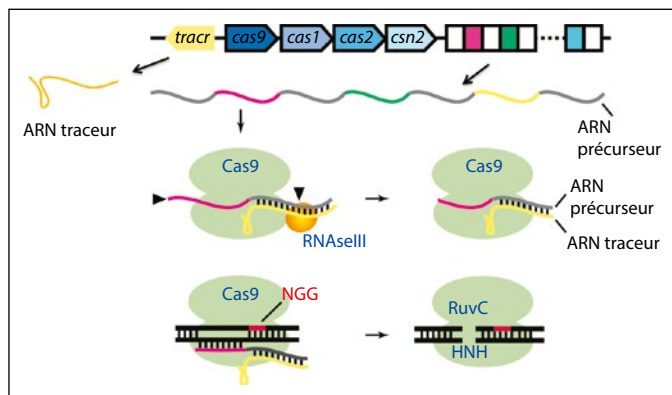


Figure 7

Expression de l'ARN précurseur et de l'ARN Traceur par les gènes CRISPR puis traitement de ces ARN par la protéine Cas et l'ARNase pour former les ARN-guides et le complexe nucléoprotéique de surveillance de la cellule.



Figure 8

Emmanuelle Charpentier et Jennifer Doudna, duo franco-américain de recherche sur les CRISPR.



Figure 9

Des modifications génétiques avec Cas9 ont été effectuées sur de nombreux animaux... dont l'humain.

Suite à ces travaux sur les bactéries (publiés en 2013), d'autres études ont porté sur ces technologies à un échelon supérieur chez la bactérie.

L'outil CRISPR-Cas9 peut par exemple être utilisé pour introduire des milliers de mutations en parallèle. Cette technique est utile pour générer des variants et sélectionner des souches améliorées, notamment à des fins de bio-production.

Une autre étude a démontré la possibilité d'utiliser CRISPR-Cas9 pour faire des modifications extrêmement importantes. Les auteurs ont utilisé ces technologies CRISPR pour aller recoder le génome, c'est-à-dire changer l'utilisation des codons chez *E. coli*. Dans le code génétique qui spécifie les triplets correspondant aux acides aminés, il y a des redondances, des acides aminés qui vont être encodés par plusieurs triplets différents. Potentiellement on pourrait tenter de libérer certains de ces triplets pour en faire autre chose, notamment incorporer des acides aminés non naturels, ce qui

peut conduire à une multitude d'applications technologiques – d'où l'intérêt de changer l'usage des codons. Ce travail demande de modifier des milliers de positions sur d'énormes morceaux d'ADN. Cela est réalisable d'une part en re-synthétisant l'ADN (on va synthétiser des fragments du génome de *E. coli* en changeant l'usage de codons), et d'autre part en introduisant ces fragments synthétiques grâce à des outils CRISPR-Cas9 dans le chromosome de la bactérie.

On le voit, toutes ces techniques, que ce soit chez la bactérie ou sur des cellules humaines ou autres, commencent toutes par reprogrammer la nucléase Cas9¹². On peut la regarder comme des ciseaux reprogrammables avec un ARN-guide, qui va permettre d'introduire une cassure à un endroit précis. En fonction de ce qui se produit avec cette cassure, on peut ensuite réaliser différentes choses.

12. Nucléase Cas9 : enzyme spécialisée pour couper l'ADN avec deux zones de coupe actives, une pour chaque brin de la double hélice.

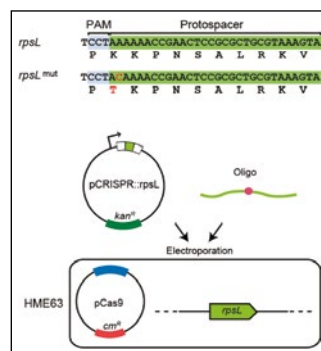


Figure 10

Insertion par électroporation du plasmide de Cas9 et de l'oligonucléotide pour modifier l'ADN d'*E. coli* de façon ciblée. *rpsL* est le gène ciblé dans cette expérience : il s'agit d'une protéine ribosomale qui, une fois mutée, donne une résistance à la Streptomycine. La séquence marquée *rpsLmut* contient la mutation introduite.

2.2. Développement d'antibiotiques intelligents

Pour procéder à des modifications précises des génomes après cassure, une première technique consiste à fournir une molécule d'ADN qui servira à réparer la cassure de manière précise par recombinaison homologue¹³. Mais une autre technique consiste simplement à réaliser cette cassure et compter sur les machineries naturelles de réparation des cellules. On peut utiliser cette dernière technique lorsqu'il existe un mécanisme de réparation – qu'on appelle « *Non Homologous End Joining* » (NHEJ) – capable de réparer ces cassures double brin, même si c'est au prix d'une certaine probabilité d'erreurs.

Cette technique est très utilisée avec CRISPR puisqu'elle permet notamment d'introduire des insertions ou de petites délétions dans les gènes et de faire des « *frame shifts* » dans le gène permettant d'éteindre les gènes cibles.

Ces deux voies de modification du génome par recombinaisons homologues et par NHEJ sont particulièrement efficaces chez les systèmes eucaryotes¹⁴. En revanche chez les bactéries, la conséquence principale d'une cassure Cas9

est en fait de tuer les cellules, qu'on ait utilisé l'une ou l'autre de ces voies d'intervention (**Figure 11**). Il est nécessaire d'introduire des recombinaisons de phage pour utiliser la recombinaison homologue de manière efficace chez les bactéries.

On s'est récemment posé la question de savoir si on pouvait tenter d'utiliser cette capacité de Cas9 à tuer les bactéries comme une stratégie antimicrobienne, une sorte d'antibiotique intelligent en reprogrammant Cas9 pour cibler spécifiquement des gènes de résistance aux antibiotiques ou des gènes de virulence, et éliminer uniquement les bactéries qui portent ces gènes sans toucher au reste du microbiome.

La **Figure 12** résume la preuve de concept. On injecte le système CRISPR-Cas9 en utilisant des bactériophages comme vecteurs¹⁵, ce qui revient à retourner complètement le système naturel. Ce système CRISPR est ensuite guidé pour cibler par exemple un gène de résistance à un antibiotique dans le chromosome de la bactérie et tuer les bactéries qui portent ce gène.

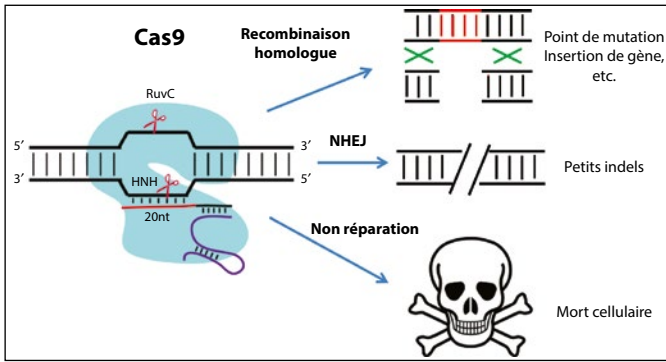
Sur la droite de la **Figure 12**, sont représentés des tapis cellulaires développés sur des boîtes de Petri de la bactérie *Staphylococcus aureus*¹⁶

13. Recombinaison homologue : type de recombinaison génétique où les séquences de nucléotides sont échangées entre des molécules d'ADN homologues.

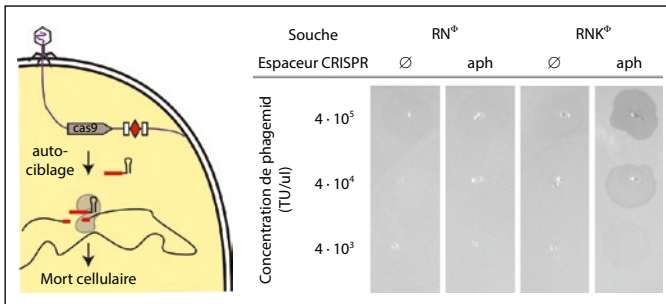
14. Eucaryote : dont les cellules possèdent un noyau structuré. Les eucaryotes rassemblent quatre grands règnes du monde du vivant : les animaux, les champignons, les plantes et les protistes.

15. Vecteur : un organisme qui ne provoque pas lui-même une maladie mais qui disperse l'infection en transportant les agents pathogènes d'un hôte à l'autre.

16. *Staphylocoque aureus* : le staphylocoque doré est responsable d'intoxications alimentaires et d'infections potentiellement mortelles.


Figure 11

Trois effets possibles de l'action d'un système CRISPR-Cas9 : recombinaison homologue, NHEJ ou mort de la cellule.


Figure 12

Sur des bactéries incapables de se réparer, les ciseaux CRISPR-Cas9 ont pour effet de tuer les bactéries (taches sur la droite).

RN : *Staphylococcus aureus* RN4220 ; RNK : variant de cette souche résistant à la kanamycine ; Aph : gène de résistance à la kanamycine [désigne dans ce cas le fait que l'espaceur CRISPR a été conçu pour cibler ce gène] ; TU : transducing units, nombre de particules de phage utilisé.

(staphylocoques dorés) et sur lesquels on a déposé une goutte de cette préparation de CRISPR vectorisée par des bactériophages. Ce qu'on observe, c'est que sur la droite il s'agit de tapis de staphylocoques qui portent un gène de résistance à la kanamycine¹⁷, et sur la toute droite c'est le système CRISPR programmé pour cibler ce gène de résistance à la kanamycine qui a été déposé. On observe qu'en

déposant la goutte sur le tapis de bactérie, on tue de manière efficace les bactéries cibles (**Figure 12**). On a en fait réalisé un antibiotique à précision ultime, puisqu'on est capable d'aller tuer les bactéries en fonction des gènes qu'elles portent.

2.3. Des antibiotiques à la précision ultime utilisant la compétition bactérienne naturelle

Un intérêt de réaliser des antibiotiques aussi spécifiques est

17. Kanamycine : antibiotique qui peut traiter une large variété d'infections.

de profiter de la compétition entre les bactéries pour occuper une niche¹⁸.

Une expérience simple illustre ce concept. On cultive une population de staphylocoques dorés, dont certains sont résistants et d'autres sensibles à la kanamycine, et on les met en compétition. Un marqueur de fluorescence GFP¹⁹ a été introduit dans les résistants pour suivre facilement les deux populations dans une co-culture. Le graphe de la **Figure 13** montre l'évolution des populations. La ligne pleine mesure juste la densité optique de la co-culture : elle donne la totalité des bactéries qui poussent dans le milieu. La ligne en

pointillés est le signal de fluorescence et indique donc la quantité des bactéries résistantes. L'expérience contrôle est figurée en bleu.

Dans une autre expérience, on traite avec de la kanamycine et donc on tue tous les sensibles ; il ne reste que les résistants et le signal fluorescent est très élevé (en orange).

Avec un antibiotique non sélectif, par exemple de la streptomycine²⁰, on tue les deux populations de bactéries : les résistants et les sensibles à la kanamycine. Cela se traduit par un délai dans la croissance, mais au bout d'un moment, cela repart car des résistants vont apparaître dans les deux populations à une fréquence équivalente (en vert). *In fine* on se retrouve donc tout de même avec la moitié de la population qui est fluorescente correspondant aux bactéries résistantes à la kanamycine, et on n'a pas réussi à éliminer les bactéries résistantes.

En revanche, si on traite avec un système CRISPR programmé spécifiquement pour cibler ce gène de résistance à la kanamycine, on observe les variations présentées par la courbe violette : à la fin de l'expérience, on n'a pas récupéré de signal de fluorescence et la ligne de pointillés est toujours autour de 0 (en violet). On a tué peut-être 99 % des bactéries résistantes pendant que les autres bactéries, les

18. Niche : ici, niche cellulaire, microenvironnement cellulaire.

19. GFP : « Green Fluorescent Protein », protéine ayant la propriété d'émettre une fluorescence de couleur verte.

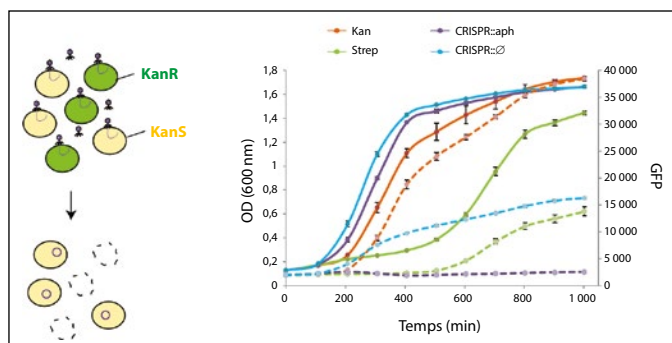


Figure 13

En bleu : expérience contrôle : co-culture de bactéries résistantes (ou pas) à la kanamycine (KanR ou KanS), au bout de 1 000 min, la population est à peu près équitablement répartie. En orange : ajout de kanamycine, toutes les bactéries KanS meurent, il ne reste plus que des KanR au bout de 1 000 min. En vert : expérience de contrôle : en présence de Streptomycine, les deux populations KanR et KanS meurent et se reconstituent, au bout de 1 000 min la population est à peu près équitablement répartie. En violet : action du système CRISPR-Cas9 sur le gène résistant des KanR ce qui éteint quasi totalement la population et peuplement de la niche par les KanS non affectées.

20. Streptomycine : antibiotique à spectres larges pouvant réagir avec les bacilles gram négatifs, avec gram positifs ou avec certaines mycobactéries.

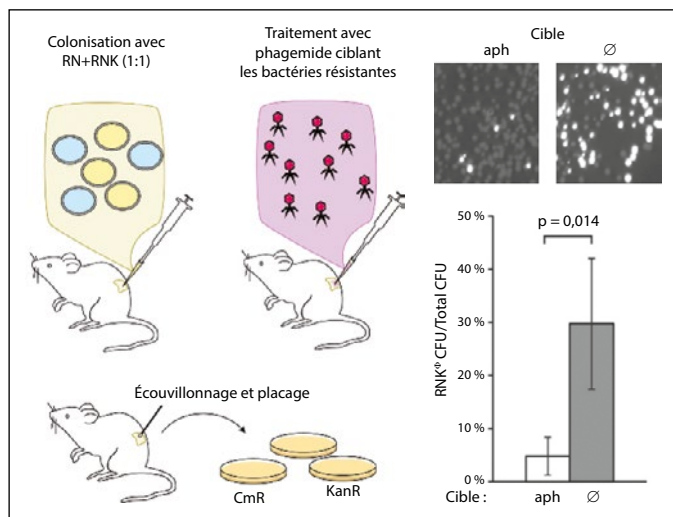


Figure 14

Expérience analogue mais sur de la peau de souris. On obtient de bons résultats avec diminution effective du nombre de bactéries résistantes (*aph*) par rapport à la situation normale (Ø).
CFU : Colony Forming Units, nombre de bactéries dans l'échantillon

sensibles, continuent à pousser dans la culture (courbe violette pleine).

On voit par cet exemple qu'on peut tirer profit de la compétition qui existe entre les souches bactériennes pour occuper une niche, on va pouvoir être plus efficace que n'importe quelle autre stratégie antibiotique pour éliminer spécifiquement une catégorie de bactérie.

Au-delà du monde des cellules, on a pu réaliser des expériences animales avec CRISPR (Cas9). L'exemple présenté met en jeu des souris. On rase leur dos et on colonise les souris avec le mélange de staphylocoques résistants et non résistants à la kanamycine. On traite avec notre préparation de CRISPR vectorisée pour cibler le gène de résistance à la kanamycine, et on voit que dans ce modèle de colonisation de la peau également, on a été capable de diminuer spécifiquement la colonisation des bactéries résistantes (Figure 14).

Ces travaux très prometteurs sont poursuivis maintenant par la start-up Eligo Bioscience, basée à l'hôpital Cochin. On continue ainsi d'avancer sur ces stratégies en vue de les amener idéalement jusqu'à des essais cliniques chez l'homme.

3 Les autres applications des CRISPR

3.1. Moduler l'expression des gènes pour comprendre les réseaux génétiques

Dans une autre classe d'applications technologiques du système CRISPR chez les bactéries, on utilise un variant de la protéine Cas9 que l'on appelle catalytique nul, ou dCas9 : il n'est plus capable de couper l'ADN (Figure 15). Cependant, il peut, guidé par ses petits ARN, s'attacher sur une séquence cible de l'ADN sans le couper ; il reste attaché suffisamment fortement

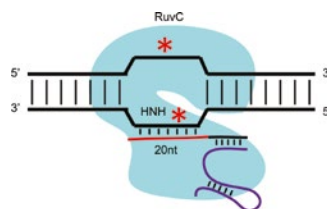


Figure 15

Variant Cas9 appelé dead Cas9 (dCas9).

pour éteindre l'expression des gènes cibles.

La **Figure 16** résume les résultats d'expériences où l'on a ciblé un rapporteur²¹ de fluorescence GFP au moyen d'ARN-guides, qui vont cibler soit le brin codant ou le brin Template, soit par l'intérieur du gène rapporteur, soit par sa région promotrice ; on voit qu'on est capable, grâce à cet outil, d'éteindre de 100 à 1 000 fois l'expression du gène cible.

Cet outil, extrêmement performant, permet de contrôler à façon les expressions des gènes chez la bactérie et maintenant aussi sur des cellules de souris, et même sur des cellules humaines.

21. Rapporteur : un gène rapporteur est un gène dont le produit (protéine) possède une caractéristique lui permettant d'être observé en laboratoire (fluorescence, activité enzymatique détectable...).

Un outil similaire permet d'activer l'expression de gènes : on utilise toujours la protéine Cas9, mais en y fusionnant un domaine capable de recruter une polymérase, promouvant ainsi l'expression des gènes cibles.

Nos travaux récents examinent s'il est possible non plus d'éteindre complètement l'expression d'un gène mais d'obtenir des niveaux intermédiaires finement régulés de leur expression. Pour ce faire, on a comparé deux approches : soit contrôler la concentration en protéine Cas9 dans la bactérie, soit introduire des mutations dans l'ARN-guide pour qu'il ne soit plus parfaitement homologue à la cible, et déstabiliser ainsi l'interaction entre Cas9 et l'ADN pour permettre un niveau intermédiaire d'expression.

Les expériences montrent que les deux stratégies donnent des résultats mais à des

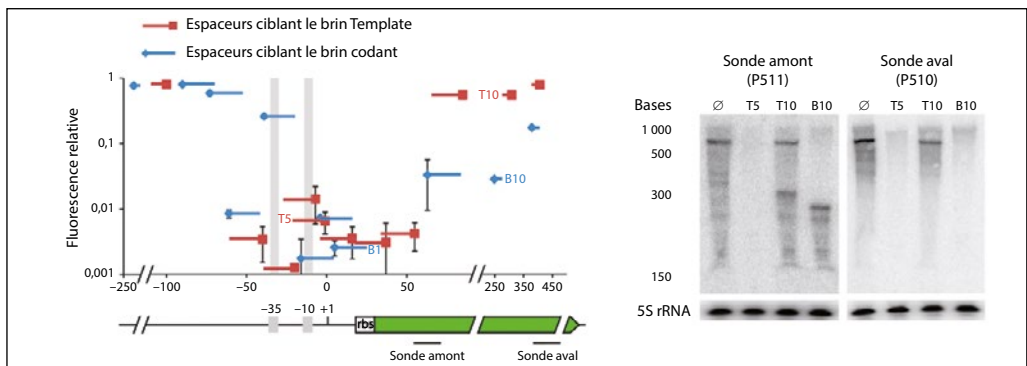


Figure 16

Répression d'un gène rapporteur fluorescent (GFP) par des ARN-guides ciblant le gène ou son promoteur dans les deux orientations possibles. Un « Northern Blot » permet d'analyser l'ARN produit lors de cette expérience et en présence de différents ARN-guides (T5, T10 ou B10). Cette expérience révèle que *dCas9* bloque la progression de l'ARN polymérase conduisant à la formation d'un ARN messager tronqué.

En abscisses : la distance en paires de bases au promoteur du gène.

degrés différents (**Figure 17**) : lorsqu'on contrôle les niveaux intermédiaires d'expression en introduisant des modifications par mutations de l'ARN, on a des niveaux d'expression stables et très peu bruités (résultats correspondant à la partie gauche de la **Figure 17**). En bas de la figure sont schématisées les distributions d'expressions obtenues au sein d'une population de bactéries ; elles sont extrêmement fines sur la gauche mais très larges sur la droite. Cette approche est très puissante et peut être multiplexée : on peut cibler plusieurs gènes simultanément.

Ces techniques devraient se révéler extrêmement utiles pour comprendre le fonctionnement des cellules, les réseaux génétiques et leurs interactions. On peut y parvenir grâce à des rapporteurs qui expriment une protéine de fluorescence rouge sur un gène et une protéine de

fluorescence verte sur un autre ; ces deux gènes seront ensuite contrôlés de manière fine afin de pouvoir explorer tout le paysage stœchiométrique d'expression de ces différents gènes.

3.2. Bloquer l'expression de gènes et séquençage pour mieux comprendre les gènes

Citons encore une application étonnante de ces outils CRISPR chez la bactérie. Il s'agit toujours d'utiliser Cas9 pour bloquer l'expression de gènes, mais on l'envisage avec des techniques à haut débit – ne pas bloquer un seul gène, mais simultanément une quantité importante de gènes. Plus précisément, il s'agit de synthétiser des morceaux d'ADN qui vont porter l'ARN-guide, qu'on va être capable d'introduire dans les plasmides et dans les bactéries. On se retrouve ainsi avec des bibliothèques de 10^5 , 10^6 ARN-guides différents pour cibler

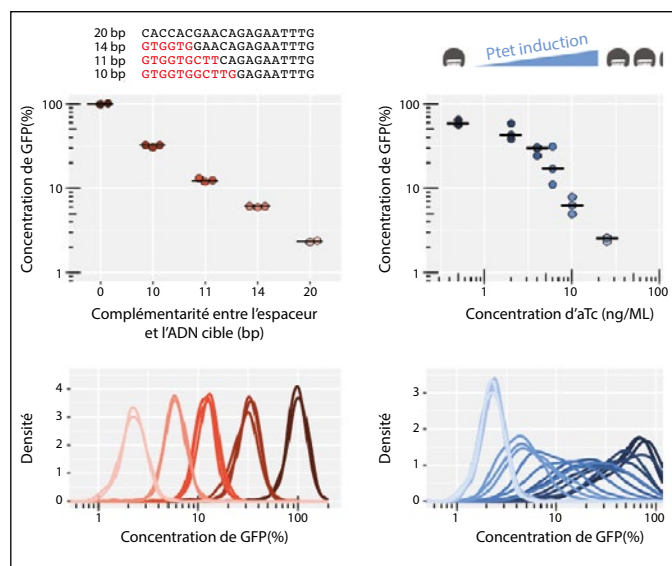
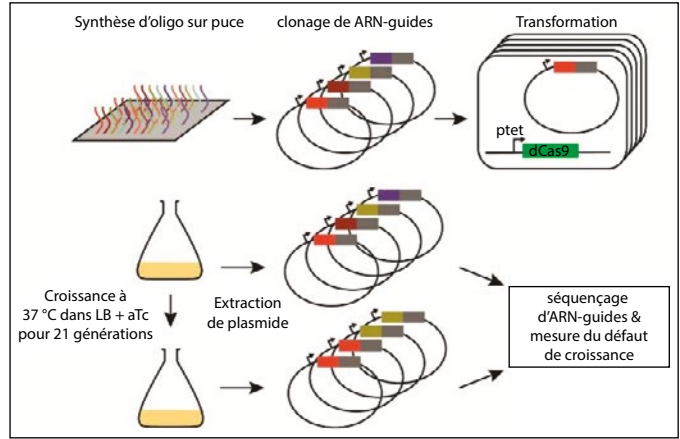


Figure 17

L'expression d'un gène rapporteur cible (GFP) peut être finement contrôlée soit en introduisant des bases non complémentaires entre l'ARN-guide et la cible (a), soit en diminuant la concentration de dCas9 dans la cellule par l'intermédiaire d'un promoteur inducible à l'androtetracycline (aTc) (b). Ces deux méthodes ont cependant des propriétés bien différentes à l'échelle de la population. On observe des populations avec une répression uniforme pour la première méthode (d), tandis que l'expression est très bruitée pour la deuxième méthode (c).

Figure 18

Principe du screening : synthèse d'oligonucléotides sur puces à ADN, clonage, transformation dCas9, culture, extraction et séquençage.
aTc = inducteur de Cas9.



un grand nombre de positions dans le génome de la bactérie [Figure 18]. On aura accès à des mesures de l'effet de chacun de ces ARN-guides dans la population en utilisant une technique de séquençage.

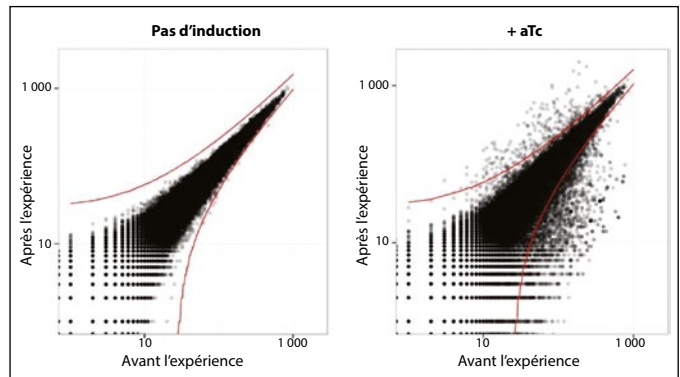
On séquence la banque d'ARN-guide à la fois au début et à la fin de l'expérience et on regarde quels sont les ARN-guides qui sont à la fois perdus lors de l'expérience et ceux qui sont enrichis pendant l'expérience. Pour ces derniers, le gène qu'ils ont éteint leur a donné un avantage alors que pour les gènes « déplétés », le gène que l'on a éteint a tué

ou ralenti fortement la croissance de la bactérie.

Sur les Figures 19 et 20, chaque point des graphes correspond à un ARN-guide différent. On a donc 10^5 guides qui vont cibler le génome d'*E. coli*, et on peut comparer l'abondance de chacun de ces guides dans la population en début d'expérience et en fin d'expérience. Sur la gauche de la figure, il s'agit d'un contrôle où l'on n'a pas exprimé dCas9 : tout reste sur la diagonale, ce qui veut dire que tout va bien (cela ne change pas entre le début et la fin de l'expérience). En revanche, lorsque l'on va

Figure 19

Expérience avant/après en présence d'aTc (inducteur de dCas9) ou pas (non induction), certains points s'écartent de la diagonale dans le cas de l'induction de dCas9, ce qui montre que les ARN-guides sont dans certains cas enrichis (montés) ou perdus (descendus). En abscisse : nombre de lectures de séquençage par la technologie Illumina. Ce nombre reflète l'abondance du guide dans la population.



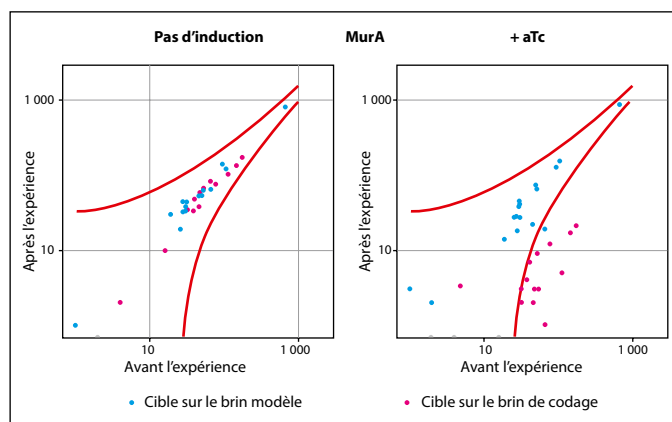


Figure 20

Même tracé pour le gène essentiel synthèse de paroi de cellule de *E. Coli* pour illustrer les ARN-guides qui bloquent l'expression de *MurA* (descendus).

induire l'expression de dCas9, on se retrouve avec beaucoup de points qui vont descendre ou qui vont monter (partie droite de la figure). Ceux qui descendent correspondent à des ARN-guides perdus au cours de l'expérience et ceux qui montent à des ARN-guides enrichis au cours de l'expérience (Figure 19).

La Figure 20 montre le même diagramme, où l'on n'a gardé que les points correspondant à des ARN-guides qui ciblent un gène essentiel

pour *E. Coli* – en l'occurrence le gène *MurA* nécessaire à la synthèse de la paroi de la cellule. Les ARN, en rouge, qui bloquent l'expression de ce gène de manière efficace, descendent tous et sont donc tous perdus au cours de l'expérience. Ce type d'approche ouvre la promesse de redécouvrir tous les gènes essentiels de l'organisme ; en variant les conditions expérimentales, on peut répondre à toute une variété de questions (Figure 20).

Le début des infinies possibilités de Cas9

Nous avons successivement rappelé que cet outil Cas9, qui a fait son apparition récemment dans le monde scientifique, est extrêmement puissant, autant pour la modification des génomes chez les eucaryotes, notamment

les hommes, avec des exemples de thérapie génique, mais également chez la bactérie.

On a montré que Cas9 peut être utilisé pour tuer les bactéries de manière très efficace. Cette propriété est potentiellement riche d'applications. Les laboratoires les étudient et développent peut-être jusqu'au stade clinique pour guérir efficacement certaines maladies infectieuses.

Enfin, nous avons indiqué les propriétés de ce variant de Cas9, qu'on appelle dCas9, qui est capable de bloquer l'expression de gènes de manière très efficace.

Cet outil va certainement permettre des progrès spectaculaires dans l'étude du fonctionnement des cellules, aussi bien des bactéries que des cellules humaines.