

ÉCOLES NORMALES SUPÉRIEURES
ÉCOLE NATIONALE DES PONTS ET CHAUSSÉES

CONCOURS D'ADMISSION SESSION 2016

FILIÈRE BCPST
COMPOSITION DE BIOLOGIE

Épreuve commune aux ENS de Cachan, Lyon, Paris et à l'ENPC

Durée : 6 heures

L'utilisation des calculatrices n'est pas autorisée pour cette épreuve.

Thème de l'épreuve :

Les relations interspécifiques

L'épreuve comporte **un sujet de synthèse** et **un sujet avec documents**. Des durées conseillées sont données :

Sujet	Pages	Durée conseillée	Intitulé
Synthèse	3	2 h 30	Importance écologique des relations interspécifiques
Documents	4 à 21	3 h 30	Le système CRISPR/Cas : Mécanisme, aspects évolutifs et applications

Il vous est fortement recommandé de bien **gérer votre temps** de composition afin de pouvoir traiter correctement les différentes parties de l'épreuve.

Lors de l'évaluation, les correcteurs et correctrices attacheront une importance particulière à :

- la justification des raisonnements ;
- la clarté et la concision des réponses ;
- la qualité et la précision des illustrations ;
- l'orthographe, la grammaire et la présentation.

L'usage des calculatrices est interdit.

SUJET DE SYNTHÈSE

Importance écologique des relations interspécifiques

En utilisant préférentiellement des exemples issus de la prairie pâturée, vous présenterez la diversité des relations interspécifiques et montrerez comment celles-ci participent au fonctionnement et à la dynamique des écosystèmes.

SUJET SUR DOCUMENTS

Le système CRISPR/Cas Mécanisme, aspects évolutifs et applications

Chez certaines espèces bactériennes, le locus dit CRISPR (pour *clustered interspaced short palindromic repeats*) est hautement polymorphe. Le locus CRISPR ([Figure 1](#)) est constitué de multiples répétitions d'une même séquence d'environ 30 nucléotides, qu'on appellera « séquence répétée ». Deux séquences répétées sont séparées par une « séquence espaceur ». Les séquences espaceurs ont chacune une longueur d'environ 35 nucléotides, mais une séquence variable, unique au sein du locus. Le locus CRISPR est adjacent à un ou plusieurs gènes dits *cas* (pour *CRISPR-associated*). On s'intéresse dans ce sujet au mécanisme dit CRISPR/Cas de résistance aux éléments génétiques étrangers.

Les résultats présentés sont issus d'expériences répétées au moins 3 fois. Dans les graphiques, on représente la moyenne des résultats, les barres d'erreur correspondant à l'intervalle de confiance à 95% de la moyenne. Les images et données brutes sont représentatives de l'ensemble des résultats obtenus.

PARTIE I

Identification du mécanisme évolutif CRISPR/Cas de résistance des procaryotes aux éléments génétiques étrangers

Dans une première série d'expériences, on utilise comme système expérimental l'espèce bactérienne *Streptococcus thermophilus* (*S. thermophilus*) et les phages (virus infectant les bactéries) ϕ 858 et ϕ 2972. Le génome de ces deux phages est constitué d'une molécule d'ADN double brin. Les séquences génomiques des deux phages sont identiques à 93%. La souche bactérienne *S. thermophilus* sauvage (WT) est sensible à ces deux phages, c'est-à-dire que l'infection par les phages aboutit à la lyse bactérienne.

On procède tout d'abord à une expérience d'évolution *in vitro*. La souche bactérienne *S. thermophilus* WT est cultivée pendant plusieurs générations en présence de l'un, ou de l'autre, ou des deux phages $\phi 858$ et $\phi 2972$. Neuf souches bactériennes mutantes résistantes aux phages sont obtenues de façon indépendante et leur locus CRISPR est analysé. On quantifie leur sensibilité à chacun des deux phages (Figure 1). La sensibilité relative d'une souche bactérienne est exprimée comme le ratio du nombre de plages de lyse obtenues après culture sur milieu gélosé en présence du phage sur le nombre de plages de lyse obtenues pour la souche bactérienne WT en présence du même phage.

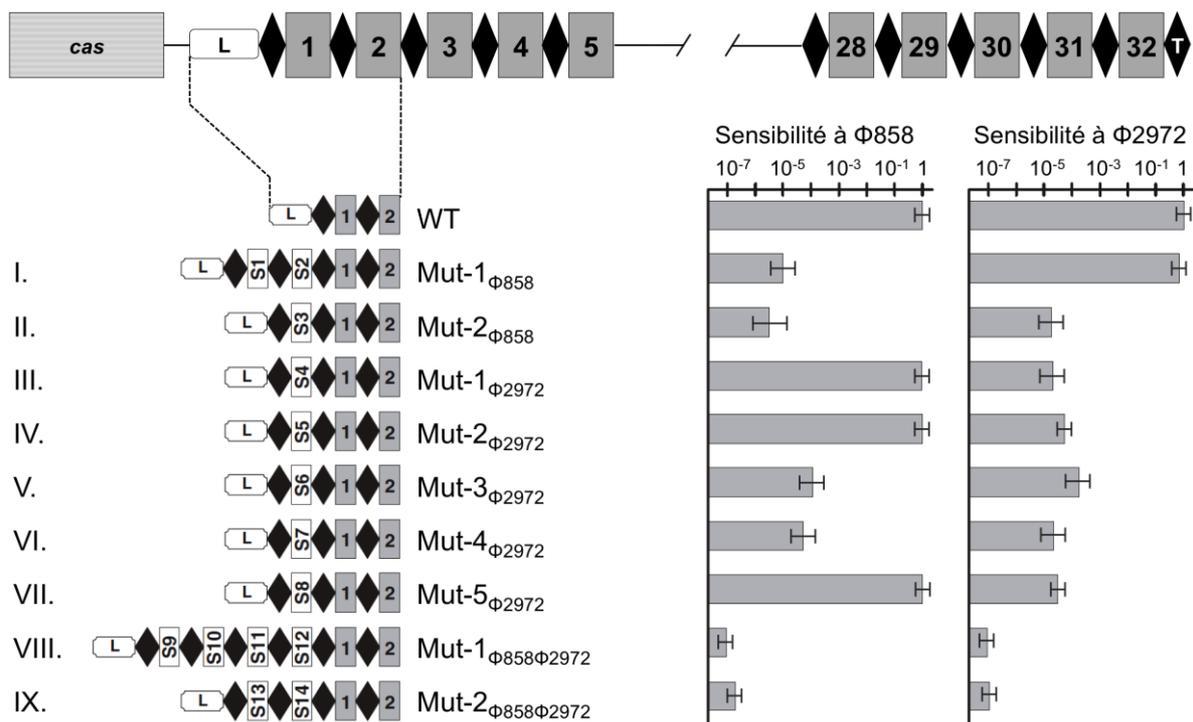


Figure 1. Evolution du locus CRISPR au cours de l'infection de *S. thermophilus* par $\phi 858$ et $\phi 2972$.

En haut, structure du locus CRISPR de *S. thermophilus* WT. La séquence répétée est symbolisée par des losanges noirs. Les séquences espaceurs sont numérotées de 1 à 32 et symbolisées par des rectangles gris. Le locus est bordé par une séquence L en amont comportant le site d'initiation de la transcription et une séquence T en aval comportant le site de terminaison de la transcription. Les gènes *cas* adjacents au locus CRISPR sont également symbolisés.

En bas, structure du locus CRISPR des souches bactériennes résistantes (numérotées I à IX) obtenues après culture en présence de $\phi 858$ (souches I et II notées Mut $_{\phi 858}$), en présence de $\phi 2972$ (souches III à VII notées Mut $_{\phi 2972}$) ou en présence des deux phages (souches VIII et IX notées Mut $_{\phi 858\phi 2972}$). Les séquences espaceurs présentes dans la souche bactérienne WT sont symbolisées par des rectangles gris. Les séquences espaceurs absentes dans la souche bactérienne WT sont numérotées de S1 à S14 et symbolisées par des rectangles blancs. Pour chaque souche bactérienne, la sensibilité relative à chacun des phages est quantifiée sur les histogrammes.

On analyse alors les séquences espaceurs S1 à S14. En particulier, on compare ces séquences aux séquences génomiques des phages $\phi 858$ et $\phi 2972$ (Figure 2).

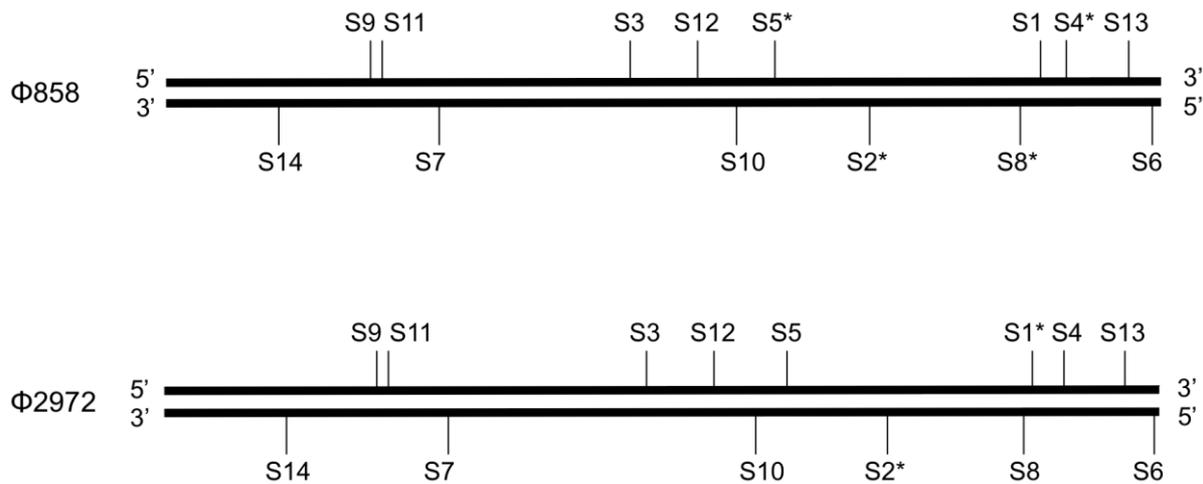


Figure 2. Carte du génome des phages $\phi 858$ et $\phi 2972$. La position des séquences de phages présentant un fort taux d'identité avec les séquences S1 à S14 est indiquée. L'absence d'astérisque indique 100% d'identité entre la séquence S et la séquence du phage. Un astérisque (*) indique l'existence de polymorphismes nucléotidiques entre la séquence S et la séquence du phage. Les séquences indiquées au-dessus du génome du phage présentent un fort taux d'identité avec le brin codant du génome du phage. Les séquences indiquées en dessous du génome du phage présentent un fort taux d'identité avec le brin non codant du génome du phage.

Question 1. Analysez la structure du locus CRISPR des souches résistantes présentées en figure 1. A l'aide de la figure 2, formulez une hypothèse expliquant l'origine des séquences S1 à S14.

Question 2. Regroupez les souches I à IX en différentes catégories en fonction de leur sensibilité à chacun des deux phages. A l'aide de la figure 2, proposez une hypothèse pouvant expliquer les profils de résistance des différentes souches.

Afin de valider ces observations, la séquence du locus CRISPR de la souche I est modifiée : on procède à la délétion des séquences répétées et espaceurs (souche X). Par ailleurs, on procède à l'insertion, au sein du locus CRISPR de la souche III, des séquences espaceurs S1 et S2 séparées par une séquence répétée (souche XI). On mesure alors la sensibilité relative des souches bactériennes ainsi obtenues à l'infection par $\phi 858$ (Figure 3).

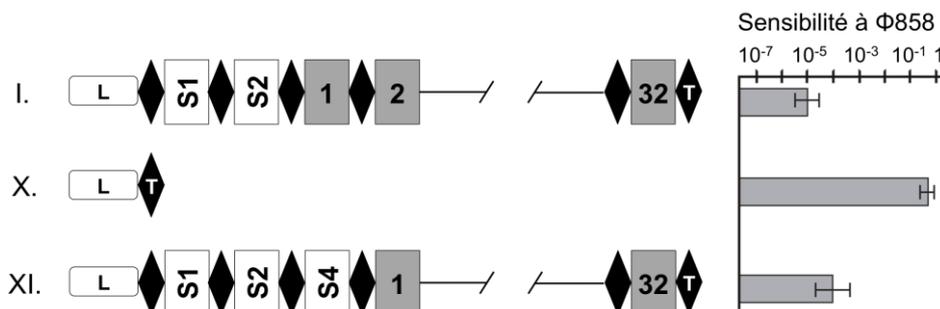


Figure 3. Modification artificielle des locus CRISPR et impact sur la sensibilité à $\phi 858$. A partir du locus CRISPR de la souche bactérienne I, on délète les séquences répétées et espaceurs. La souche bactérienne résultante est nommée X. On insère par ailleurs les séquences espaceurs S1 et S2 au sein du locus CRISPR de la souche bactérienne III. La souche bactérienne résultante est nommée XI. La sensibilité relative des souches bactériennes I, X et XI au phage $\phi 858$ est quantifiée dans l'histogramme.

Question 3. Analysez et interprétez la figure 3.

Dans une deuxième série d'expériences, on utilise comme système expérimental l'espèce bactérienne *Staphylococcus epidermidis* (*S. epidermidis*). Cette bactérie peut acquérir des plasmides étrangers par transfert horizontal et en particulier par conjugaison. Le génome de la souche *S. epidermidis* sauvage (WT) comporte un locus CRISPR qui inclut une séquence espaceur identique à une séquence du gène *nes* porté par le plasmide pNes(wt), un plasmide qui peut être transmis par conjugaison.

On génère une souche bactérienne mutante (*S. epidermidis* mut) dans laquelle le locus CRISPR est délété. On génère également un plasmide pNes(mut) dans lequel des mutations silencieuses ont été introduites dans le gène *nes*. On analyse alors l'efficacité de l'acquisition de chacun des deux plasmides pNes(wt) et pNes(mut) par conjugaison chez les deux souches de bactéries *S. epidermidis* WT et *S. epidermidis* mut (Table 1).

	Souche bactérienne			
	<i>S. epidermidis</i> WT		<i>S. epidermidis</i> mut	
Plasmide	pNes(wt)	pNes(mut)	pNes(wt)	pNes(mut)
Efficacité de conjugaison	0	$2,3 \times 10^{-4}$	$2,8 \times 10^{-5}$	$8,2 \times 10^{-5}$

Table 1. Efficacité d'acquisition des deux plasmides pNes(wt) et pNes(mut) par conjugaison chez les deux souches bactériennes *S. epidermidis* WT et *S. epidermidis* mut. L'efficacité de conjugaison est calculée comme le rapport du nombre de bactéries transconjugantes (c'est-à-dire ayant acquis le plasmide) sur le nombre total de bactéries dans chaque condition.

Question 4. Analysez et interprétez la table 1.

Pour préciser le mécanisme CRISPR/Cas, on utilise le plasmide pNes(wt) dans lequel on a inséré un intron auto-épissable au sein de la séquence de *nes* identique à la séquence espaceur du locus CRISPR. Ce plasmide modifié est appelé pNes(intron).

On mesure alors l'efficacité de l'acquisition de pNes ou de pNes(intron) par conjugaison chez les souches bactériennes *S. epidermidis* WT et *S. epidermidis* mut (Table 2).

Plasmide	Souche bactérienne			
	<i>S. epidermidis</i> WT		<i>S. epidermidis</i> mut	
	pNes(wt)	pNes(intron)	pNes(wt)	pNes(intron)
Efficacité de conjugaison	0	$5,6 \times 10^{-6}$	$2,0 \times 10^{-5}$	$7,9 \times 10^{-6}$

Table 2. Efficacité d'acquisition des deux plasmides pNes et pNes(intron) par conjugaison chez les deux souches *S. epidermidis* WT et *S. epidermidis* mut. L'efficacité de conjugaison est calculée comme le rapport du nombre de bactéries transconjugantes sur le nombre total de bactéries.

Question 5. En quoi le protocole utilisé en table 2 permet-il de déterminer la cible du mécanisme CRISPR/Cas ? Analysez et interprétez la table 2.

On infecte ensuite par le phage $\phi 2972$ la souche bactérienne *S. thermophilus* WT (sensible à l'infection) et la souche bactérienne mutante VI (résistante à l'infection ; voir Figure 1). À différents temps post-infection, l'ADN est extrait des bactéries infectées, digéré à l'aide de l'enzyme de restriction *Hpa*II et les différents fragments obtenus sont analysés par Southern blot à l'aide de deux sondes oligonucléotidiques radiomarquées de séquence complémentaire au génome du phage (Figure 4).

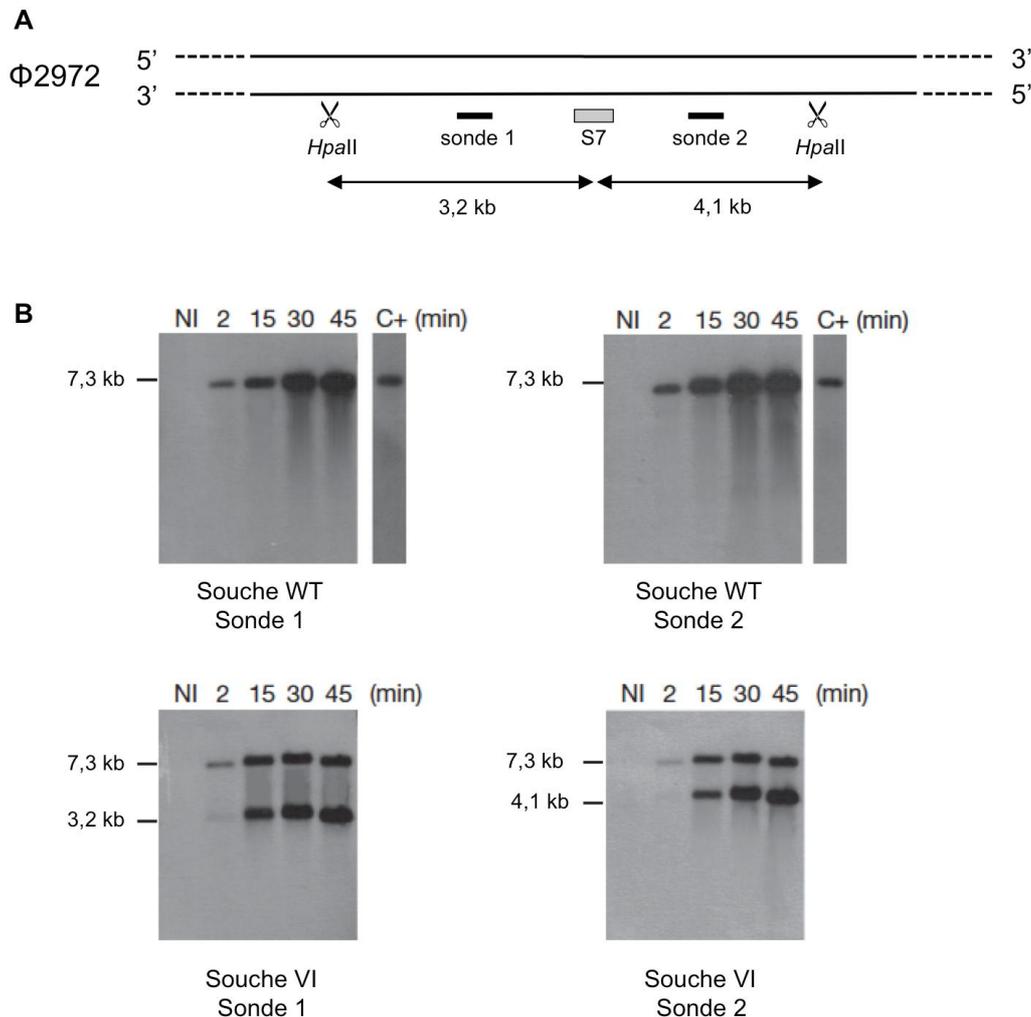


Figure 4. Analyse de l'ADN extrait des souches bactériennes *S. thermophilus* WT et mutante VI après infection par le phage $\phi 2972$.

A. Carte partielle du génome du phage $\phi 2972$ indiquant la position des sites de restriction par l'enzyme *Hpa*II, de la séquence identique à la séquence espaceur S7 et des séquences complémentaires aux sondes 1 et 2 utilisées en **B**.

B. Les souches bactériennes WT et mutantes VI sont infectées par le phage $\phi 2972$, ou non infectées (NI). Après 2, 15, 30 et 45 minutes d'infection, l'ADN est extrait, digéré à l'aide de l'enzyme de restriction *Hpa*II et les différents fragments obtenus sont analysés par Southern blot à l'aide des sondes oligonucléotidiques 1 et 2 de séquence complémentaire au génome du phage. C+, contrôle positif : l'ADN purifié à partir du phage est digéré par *Hpa*II. La même quantité d'ADN a été déposée dans chaque puits. kb : kilobases.

Question 6.

a) Analysez et interprétez la figure 4.

b) Résumez alors les informations obtenues dans cette première partie sur le mécanisme CRISPR/Cas.

PARTIE II

Modalités moléculaires du mécanisme CRISPR/Cas

Afin de caractériser le mécanisme moléculaire à l'origine du phénotype de résistance observé précédemment, on étudie le système CRISPR/Cas de l'espèce bactérienne *Escherichia coli* (*E. coli*). Le locus CRISPR de cette bactérie comporte des séquences espaceurs dont plusieurs sont identiques à des séquences du phage λ . La souche bactérienne sauvage (notée WT) comporte huit gènes *cas* adjacents au locus CRISPR, dont on cherche à définir la fonction : les cinq gènes *casA*, *B*, *C*, *D* et *E* dont les produits protéiques forment un complexe appelé Cascade, ainsi que les gènes *cas1*, *cas2* et *cas3*.

On utilise une souche bactérienne mutante d'*E. coli* (*E. coli mut*) dans laquelle les gènes *cas* sont délétés. On introduit dans cette souche *E. coli mut* les 5 gènes *casA*, *B*, *C*, *D* et *E* et/ou les gènes *cas1*, *cas2* ou *cas3* d'*E. coli* WT. Pour chacune des souches bactériennes obtenues, on quantifie la sensibilité relative au phage λ , exprimée comme le ratio du nombre de plages de lyse obtenues après culture sur milieu gélosé en présence du phage sur le nombre de plages de lyse obtenues pour la souche bactérienne *E. coli mut* (Figure 5).

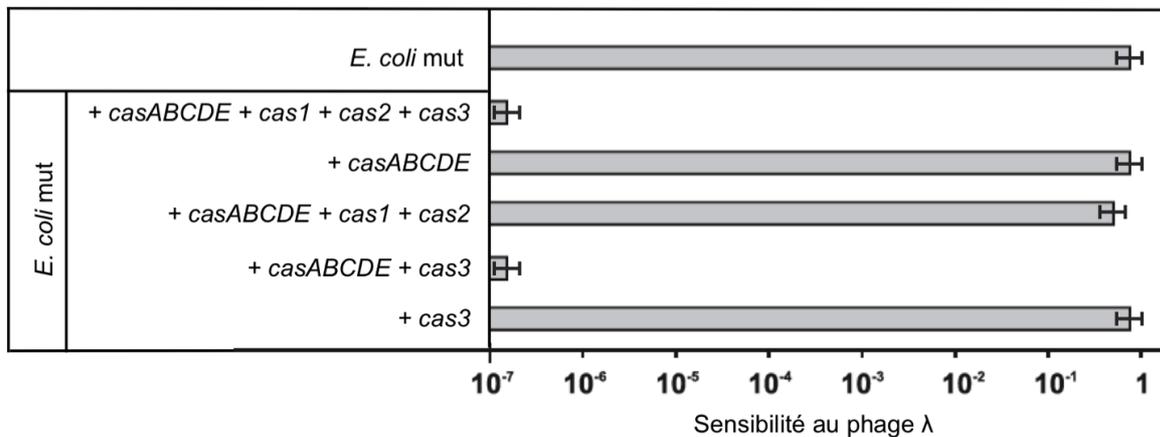


Figure 5. Caractérisation fonctionnelle des gènes *cas* chez *E. coli*. On introduit dans la souche bactérienne *E. coli mut* les 5 gènes *casA*, *B*, *C*, *D* et *E* (*casABCDE*) et/ou les gènes *cas1*, *cas2* ou *cas3* d'*E. coli* WT, comme indiqué. Pour chaque souche bactérienne ainsi obtenue, la sensibilité relative au phage λ est quantifiée sur l'histogramme.

Question 7. Analysez et interprétez la figure 5.

Question 8.

- a) Analysez la figure 6A. Proposez une ou plusieurs hypothèses rendant compte de la détection de plusieurs espèces d'ARN de taille différente.
- b) Justifiez les variations de la séquence non grisée des ARNcr, située entre les séquences 5' et 3' conservées (figure 6B, en bas).
- c) D'après la figure 6B et la taille des molécules d'ARN détectées en figure 6A, précisez alors l'hypothèse formulée en a).

Question 9. D'après vos connaissances, à quel type d'ARN décrit chez les eucaryotes pourraient s'apparenter les ARNcr ? Justifiez.

On introduit alors dans la souche *E. coli* mut les 5 gènes *casA*, *B*, *C*, *D* et *E* ou seulement 4 d'entre eux selon les différentes combinaisons possibles. L'ensemble des molécules d'ARN est extrait et analysé par northern blot à l'aide d'une sonde oligonucléotidique radiomarquée de séquence complémentaire à l'une des séquences espaceurs du locus CRISPR (Figure 7).

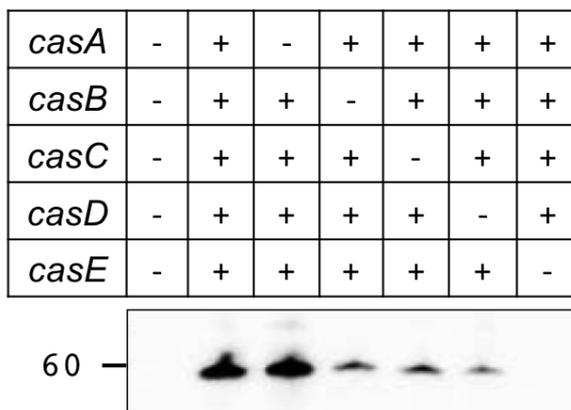


Figure 7. Caractérisation du complexe Cascade chez *E. coli*. Les molécules d'ARN sont extraites de la souche *E. coli* mut exprimant ou non 4 ou 5 des gènes *casA*, *B*, *C*, *D* et *E* d'*E. coli* WT, comme indiqué. Les molécules d'ARN sont alors séparées par électrophorèse puis transférées sur membrane. La membrane est révélée à l'aide d'une sonde oligonucléotidique radiomarquée de séquence complémentaire à l'une des séquences espaceurs du locus CRISPR. A gauche est indiquée la taille en nucléotides. La même quantité d'ARN a été déposée dans chaque puits.

Question 10. Analysez et interprétez la figure 7.

On purifie alors le complexe Cascade et les molécules qui lui sont associées à partir de bactéries *E. coli* WT infectées par le phage λ . Après purification, la fraction protéique est éliminée et la fraction non protéique subit un traitement à la RNase A, ou à la DNase I, une enzyme hydrolysant spécifiquement l'ARN, ou à la DNase I, une enzyme hydrolysant spécifiquement l'ADN. La fraction ainsi obtenue est analysée par électrophorèse (Figure 8A).

Par ailleurs, un fragment radiomarqué d'ADN issu du phage λ (ADN cible), ou d'ADN de séquence aléatoire (ADN non-cible), est incubé seul ou en présence du complexe Cascade purifié à partir de bactéries *E. coli* WT non infectées. A l'issue de l'incubation, l'ADN radiomarqué est analysé par électrophorèse (Figure 8B). Notez que les électrophorèses des acides nucléiques sont réalisées dans des conditions non dénaturantes.

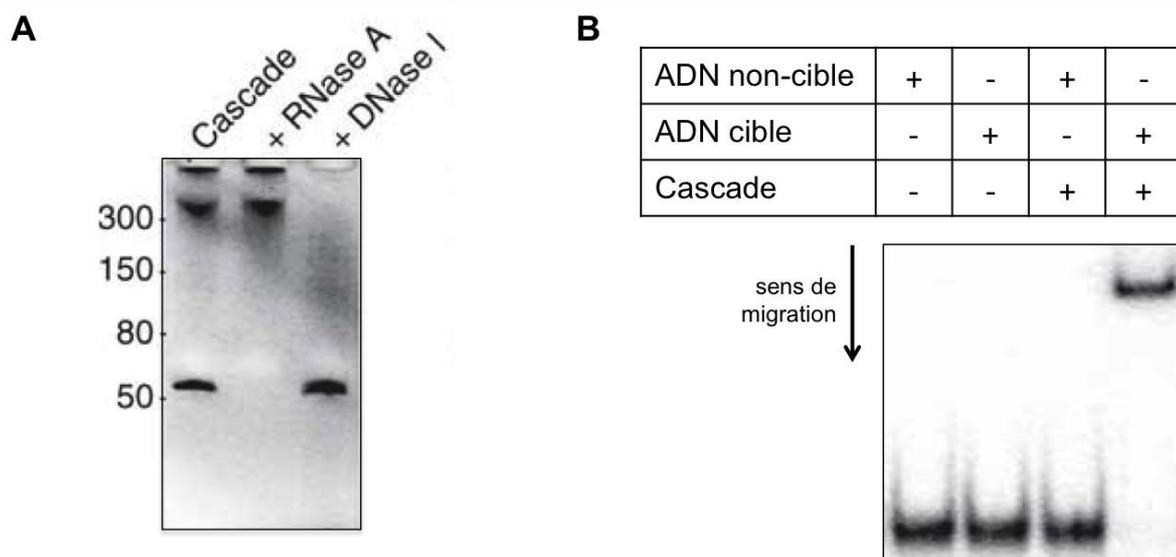


Figure 8. Identification des molécules associées au complexe Cascade.

A. Le complexe Cascade et les molécules associées sont purifiés à partir de bactéries *E. coli* WT infectées par le phage λ . Après élimination de la fraction protéique, les échantillons sont traités à la RNase A ou à la DNase I ou non traités (Cascade), puis analysés par électrophorèse en présence d'un intercalant des acides nucléiques fluorescent, et visualisés sous lumière UV. A gauche est indiquée la taille en nucléotides.

B. L'ADN cible ou non-cible radiomarqué est incubé seul ou en présence du complexe Cascade purifié à partir de bactéries *E. coli* WT non infectées, puis l'ADN radiomarqué est analysé par électrophorèse suivie d'une autoradiographie. Le sens de migration de l'électrophorèse est indiqué.

Question 11.

a) Analysez et interprétez la figure 8A.

b) Que cherche-t-on à tester dans l'expérience décrite en figure 8B ? Justifiez brièvement.

c) Analysez et interprétez la figure 8B.

On s'intéresse alors au produit du gène *cas3*. Afin de visualiser l'implication de Cas3 dans le mécanisme CRISPR/Cas, on réalise une expérience de complémentarité bimoléculaire de fluorescence. Cette technique repose sur la capacité que possèdent deux fragments d'un fluorochrome, non fluorescents lorsqu'ils sont isolés, de présenter de nouveau leur propriété de fluorescence lorsqu'ils sont en contact étroit. L'un des fragments de fluorochrome est fusionné à la protéine Cas3, et l'autre à la protéine CasA, et ces protéines sont exprimées chez *E. coli* WT. Les bactéries sont alors cultivées en absence ou en présence de phage λ et analysées par microscopie à épifluorescence (Figure 9).

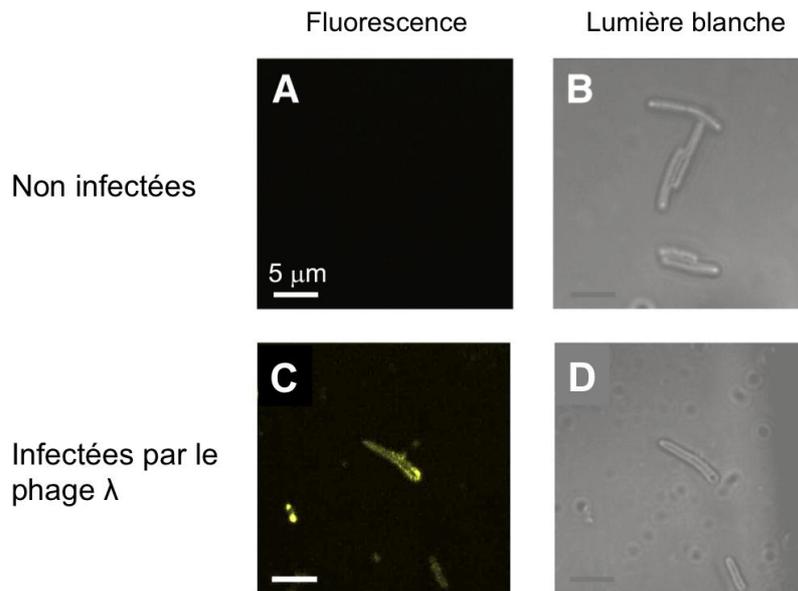


Figure 9. Caractérisation de l'implication de Cas3 par complémentarité bimoléculaire de fluorescence. Les protéines Cas3 et CasA sont fusionnées à deux fragments d'un fluorochrome puis exprimées chez des bactéries *E. coli* WT. Les bactéries sont alors infectées ou non par le phage λ puis analysées par microscopie à épifluorescence (**A** et **C**). Le même champ visualisé par microscopie en lumière blanche est présenté (**B** et **D**).

Question 12. Analysez et interprétez la figure 9.

Question 13. D'après l'ensemble des données des parties I et II, proposez des hypothèses quant au rôle de Cas3 dans le mécanisme CRISPR/Cas. Vous préciserez les étapes du mécanisme CRISPR/Cas mises en évidence jusqu'à présent.

PARTIE III

Aspects évolutifs du mécanisme CRISPR/Cas

Question 14. *Par une argumentation rigoureuse, vous explicitez les avantages et/ou désavantages évolutifs qu'un mécanisme tel que CRISPR/Cas peut présenter pour l'organisme procaryote.*

Dans le but de détecter des événements rares, on répète l'expérience de conjugaison chez *S. epidermidis* WT avec le plasmide pNes(wt), mais en augmentant le nombre de bactéries utilisées. A l'issue de l'expérience, les rares colonies transconjugantes sont isolées et on analyse leur locus CRISPR par PCR (Figure 10).

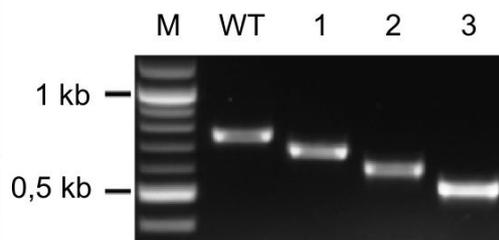


Figure 10. PCR du locus CRISPR de trois colonies bactériennes transconjugantes 1, 2 et 3, ainsi que de la souche bactérienne d'origine *S. epidermidis* WT. L'ADN de chaque souche bactérienne est purifié et amplifié par PCR en utilisant des amorces situées de part et d'autre du locus CRISPR. Les produits d'amplification sont séparés par électrophorèse en présence d'un intercalant de l'ADN fluorescent et visualisés sous lumière UV. M : marqueur de taille. kb : kilobases.

Question 15. *Proposez trois événements évolutifs pouvant chacun rendre compte du fait que des bactéries transconjugantes sont détectables malgré le système CRISPR/Cas présent chez la souche bactérienne d'origine.*

Question 16. *Analysez et interprétez la figure 10.*

Afin de déterminer si le génotype des colonies transconjugantes pré-existe à l'expérience de conjugaison (hypothèse 1) ou s'il apparaît de façon inductible lors du transfert de plasmide (hypothèse 2), on réalise une expérience dite « de fluctuation » (Table 3). On compare la variance du nombre de transconjugants obtenus dans 10 expériences de conjugaison réalisées :

- soit à partir de 10 cultures indépendantes de *S. epidermidis* WT (chacune obtenue à partir d'une colonie isolée indépendante) (protocole A) ;
- soit à partir de 10 aliquots de la même culture de *S. epidermidis* WT (protocole B).

	10 cultures indépendantes (protocole A)	10 aliquots de la même culture (protocole B)
Variance	55488	581

Table 3. Variances obtenues suite à l'expérience de fluctuation.

Question 17. Pour chacune des hypothèses 1 et 2, quelle prédiction pouvez-vous faire quant à la comparaison des variances obtenues par chacun des deux protocoles A et B ? Justifiez brièvement. Que permet de conclure la table 3 ?

On cherche alors à comparer la valeur sélective (« fitness ») associée au génotype des souches transconjugantes 1, 2 et 3 caractérisées dans la Figure 10 à celle associée au génotype de la souche d'origine *S. epidermidis* WT. Des expériences de compétition sont réalisées selon le protocole détaillé ci-dessous.

On co-cultive pendant 5 jours, à un ratio initial de 1:1, la souche bactérienne *S. epidermidis* WT avec chacune des souches transconjugantes 1, 2 et 3. On réalise également une expérience témoin dans laquelle on co-cultive la souche *S. epidermidis* WT et une souche *S. epidermidis* WT ayant acquis le plasmide pNes(mut) par conjugaison (voir Table 1). Chaque jour, la fréquence du génotype transconjugant est mesurée dans la co-culture (Figure 11).

Question 18. Soit *A* le génotype WT et *B* le génotype transconjugant. On fait l'hypothèse que la valeur sélective du génotype *A* est plus élevée que celle du génotype *B*. On considère que les populations bactériennes se multiplient de façon discrète et asexuée.

Soient les paramètres suivants (voir Table 4) :

- p_t la fréquence du génotype *A* ;
- q_t la fréquence du génotype *B* ;
- ω_B la valeur sélective relative du génotype *B* ;
- $\bar{\omega}$ la valeur sélective moyenne de la population à la génération $t - 1$, soit $\bar{\omega} = p_{t-1} + q_{t-1} \cdot \omega_B$.

	Génotype	
	A	B
Génération $t - 1$		
Fréquence avant sélection	p_{t-1}	q_{t-1}
Valeur sélective relative	1	ω_B
Génération t		
Fréquence avant sélection	p_t	q_t

Table 4. Modèle de sélection chez une population haploïde.

a) Exprimez p_t et q_t en fonction de p_{t-1} , q_{t-1} et ω_B .

b) En exprimant Δp en fonction de p_{t-1} , q_{t-1} et ω_B , montrez alors que le génotype *A* doit envahir la population dans ces conditions.

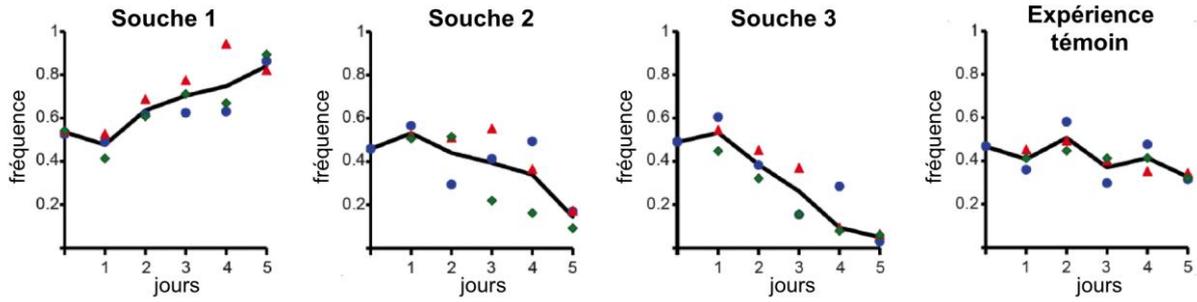


Figure 11. Fréquence du génotype transconjugant au cours du temps lors du test de compétition entre la souche bactérienne *S. epidermidis* WT et chacune des trois souches transconjugantes 1, 2 et 3. La fréquence du génotype transconjugant est indiquée en ordonnée. Le résultat de l'expérience témoin de compétition entre la souche *S. epidermidis* WT et la souche *S. epidermidis* WT ayant acquis le plasmide pNes(mut) par conjugaison est présenté à droite. Chaque série de symboles de couleurs différentes correspond à un réplicat indépendant de l'expérience. La ligne noire représente la moyenne.

Question 19.

a) Quel est l'intérêt de l'expérience témoin ?

b) Analysez la figure 11. D'après la question 18 c), qu'en déduisez-vous quant à la valeur sélective du génotype transconjugant par rapport à celle du génotype WT dans cette série d'expériences ?

c) Discutez alors la valeur sélective du système CRISPR/Cas en fonction des conditions environnementales.

PARTIE IV

Application du mécanisme CRISPR/Cas pour l'édition de gène *in vivo* chez les eucaryotes

Depuis 2012, le système CRISPR/Cas est de plus en plus exploité pour éditer les gènes eucaryotes (c'est-à-dire modifier leur séquence) *in vivo*. On utilise un système dérivé du système CRISPR/Cas de l'espèce bactérienne *Streptococcus pyogenes* et comportant deux éléments :

- la protéine Cas9 de *Streptococcus pyogenes* (SpCas9), qui est suffisante pour reconstituer un système CRISPR/Cas fonctionnel ;
- un ARNcr artificiel dont la séquence dépend du gène cible à éditer.

Dans la série d'expériences présentée ci-dessous, on souhaite éditer le gène *Mecp2* chez la souris. *Mecp2* est exprimé spécifiquement dans certains neurones et code une protéine nucléaire qui joue un rôle important dans la pathogenèse d'un syndrome neurologique appelé syndrome de Rett.

On injecte dans le cerveau deux vecteurs de transgénèse. Le transgène contenu dans le premier vecteur (vecteur SpCas9) comporte un promoteur permettant l'expression de SpCas9 suivi d'un site de poly-adénylation (pA, [Figure 12A](#)). La séquence de SpCas9 est fusionnée à une séquence de localisation nucléaire (SLN). Le transgène contenu dans l'autre vecteur (vecteur ARNcr) comporte un promoteur ubiquiste permettant la transcription de l'ARNcr ciblant *Mecp2*, et un promoteur permettant l'expression de la protéine fluorescente GFP suivi d'un site de poly-adénylation ([Figure 12B](#)).

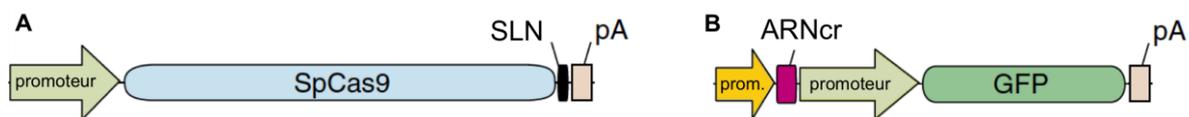


Figure 12. Schéma des transgènes véhiculés par les vecteurs SpCas9 (A) et ARNcr (B). SLN, signal de localisation nucléaire. pA, site de poly-adénylation. Prom., promoteur ubiquiste.

Question 20.

a) Quel est l'intérêt de fusionner une séquence de localisation nucléaire à la séquence codant SpCas9 ?

b) Quel est l'intérêt d'inclure la séquence codant la GFP dans l'un des transgènes ?

c) Quel promoteur utiliseriez-vous pour diriger l'expression de SpCas9 et de la GFP dans les cellules d'intérêt spécifiquement ?

Dans une première expérience, le cerveau des souris est prélevé deux semaines après l'injection des vecteurs et les cellules exprimant la GFP sont isolées. Le gène *Mecp2* est alors séquencé dans chaque cellule exprimant la GFP et les séquences obtenues sont comparées à la séquence sauvage du gène : 68% des cellules exprimant la GFP comportent une séquence différant de la séquence sauvage. Les différences sont toutes situées dans la même portion de l'exon 3 du gène ([Figure 13](#)).

Séquence sauvage	CCATTCTGCAGAGCCAGCAGAGG
	CCATTCTGCAGAGCCAGAGGCAGG
	CCATTCTGCAGAGCCCAGAGG
Séquences représentatives obtenues dans les cellules GFP ⁺	CCATTCTGCAGAGCCAGAGAGG
	CCATTCTGCAGAGCAGAGG
	CCATTCTGCAGAGCCCCAGAGG
	CCATTCTGCAGAGCCAGGAGGCCAGAGG
	CCATTCTGCAGAGCCAGAAGAGG

Figure 13. Séquence sauvage de *Mecp2* et séquences représentatives obtenues dans les cellules exprimant la GFP (GFP⁺). Seule la portion de l'exon 3 dans laquelle sont situées les différences de séquence est représentée.

Dans une deuxième série d'expériences, on injecte le vecteur SpCas9 dans deux lots de souris. Dans le premier lot de souris, on injecte également le vecteur ARNcr ciblant *Mecp2*, tandis que dans le deuxième lot de souris, on injecte un vecteur ARNcr témoin, codant un ARNcr de séquence aléatoire. Le cerveau des souris est prélevé deux semaines après l'injection des vecteurs et analysé par immunofluorescence et par western blot à l'aide d'un anticorps dirigé contre MeCP2 ([Figure 14](#)).

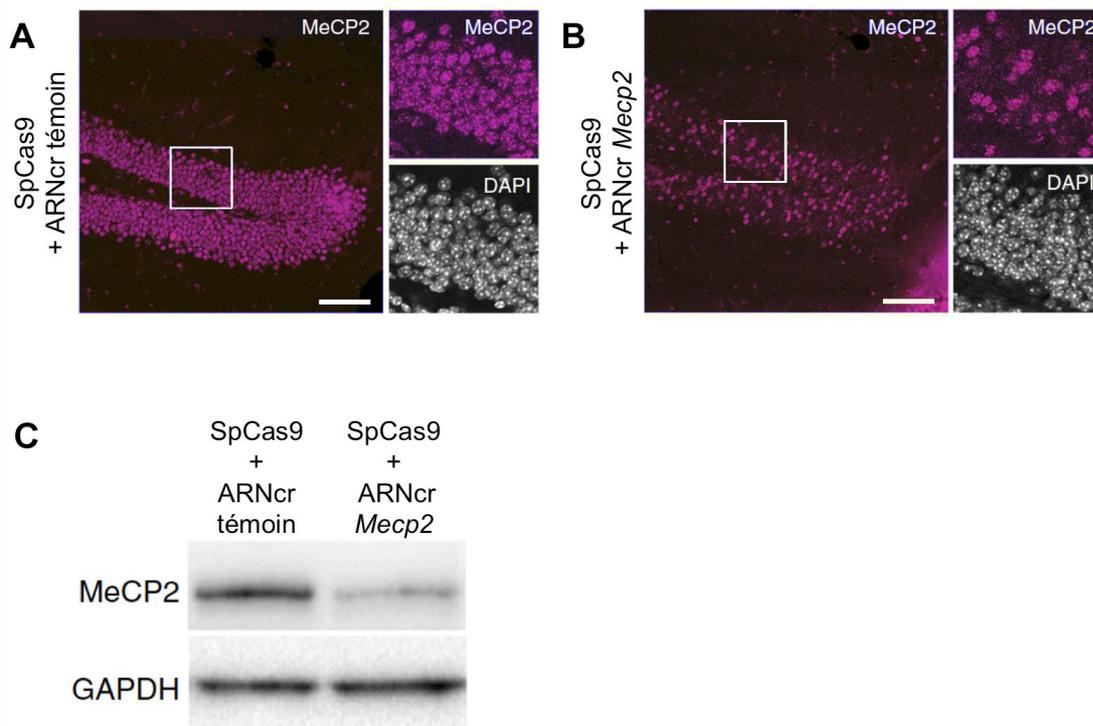


Figure 14. Efficacité de l'édition de *Mecp2* in vivo.

A et B. Des coupes histologiques de cerveau de souris chez lesquelles on a injecté le vecteur SpCas9 et le vecteur ARNcr témoin (**A**) ou ciblant *Mecp2* (ARNcr *Mecp2*, **B**) sont marquées à l'aide d'un anticorps fluorescent dirigé contre MeCP2, et à l'aide d'un intercalant fluorescent de l'ADN, le DAPI. La fluorescence correspondant au marquage de MeCP2 et au DAPI est alors visualisée à l'aide d'un microscope à épifluorescence. Les clichés de droite correspondent à des agrandissements des régions encadrées. Barre d'échelle : 150 µm.

C. Les protéines extraites du cerveau de souris chez lesquelles on a injecté le vecteur SpCas9 et le vecteur ARNcr témoin ou ciblant *Mecp2* (ARNcr *Mecp2*) sont analysées par western blot à l'aide de deux anticorps dirigés contre MeCP2 et la GAPDH, respectivement. La GAPDH est une protéine ubiquiste dont l'expression ne varie pas.

Question 21. Analysez les figures 13 et 14. Proposez des hypothèses reliant les observations de la figure 13 au mécanisme d'action du système CRISPR/Cas décrit dans les parties I et II.

Question 22. En comparaison de la technique classique de génération de souris knock-out, citez un avantage de l'exploitation du mécanisme procaryote CRISPR/Cas dans l'étude de la fonction des gènes eucaryotes.

Question 23. D'après les parties I à IV de l'exercice, proposez un schéma bilan résumant les propriétés fonctionnelles, évolutives et appliquées du mécanisme CRISPR/Cas.

Fin de l'épreuve.