

BANQUE BCPST Inter-ENS/ENPC – Session 2016

**ENS (PARIS) – ENS DE CACHAN – ENS DE LYON
ECOLE NATIONALE DES PONTS ET CHAUSSEES**

RAPPORT DE L'ÉPREUVE ÉCRITE DE BIOLOGIE

Coefficients (en pourcentage du total d'admission) :

- ENS (Paris) : Option biologie : 4,9 %, Option ST : 2,8 %
- ENS de Cachan : 12,3 %
- ENS de Lyon : Option biologie : 13,2 %, Option ST : 6,6 %
- ENPC : 5,0 %

MEMBRES DU JURY : G. BARTHOLE, A. BESSIS, J.-B. BOULE, B. DESPRES, E. GUILLAUME, C. JOURNO, V. MIRABET, P. RIALLAND-LEFEVRE, E. THIERRY.

L'épreuve écrite de Biologie de la session 2016, d'une durée totale de 6 h, comportait un sujet de synthèse d'une durée conseillée de 2 h 30 ainsi qu'un sujet d'analyse de documents d'une durée conseillée de 3 h 30, le sujet sur documents étant lui-même organisé en 4 sous-parties. Le thème commun à l'épreuve était les relations interspécifiques.

Les candidats ont globalement su faire preuve d'une bonne gestion de leur temps de composition puisque les notes obtenues dans chacune des parties étaient en moyenne équivalentes.

Sujet de synthèse : Importance écologique des relations interspécifiques

L'épreuve de synthèse évalue les compétences de communication scientifique qui figurent au programme de BCPST : **construire une argumentation scientifique** et **organiser une production écrite**. Elle a fait appel cette année à des notions nouvellement apparues dans le programme, au sein de la **partie III** (Populations, écosystèmes, biosphère). Afin de permettre aux candidats de conserver plus aisément une cohérence dans la progression de leur démonstration, et en accord avec le programme, il leur était demandé de traiter le sujet préférentiellement à partir de l'exemple de la prairie pâturée. Il était également précisé que par « importance écologique », on entendait l'implication dans le fonctionnement et la dynamique des écosystèmes.

Ce sujet demandait aux candidats de mobiliser des connaissances présentées dans les points III-A (Les populations et leur dynamique) et III-B (Les écosystèmes, leur structure et leur fonctionnement) du programme. La difficulté ne résidait donc pas dans la sélection de connaissances dispersées dans le programme mais bien dans leur articulation suivant le fil

conducteur imposé des relations interspécifiques. La **construction d'une démarche logique, cohérente et justifiée** a été valorisée, au contraire de la simple juxtaposition sans discernement de divers éléments de cours. Dans cette longue épreuve, il est tentant de se lancer le plus rapidement possible dans l'écriture afin d'exposer un grand nombre de notions. Néanmoins, cette stratégie ne répond pas aux compétences de communication scientifique attendues et produit le plus souvent des exposés qui se réduisent à une accumulation déstructurée de faits qui ne sont pas toujours en lien avec le sujet et qui ne sont pas intégrés dans une démarche hypothético-déductive. Ainsi, de nombreuses copies ont présenté des hors sujets, laissant supposer que trop peu de candidats ont pris suffisamment de temps pour délimiter le sujet et construire l'articulation de leur exposé. Nous insistons également sur le fait que l'écologie ne se résume pas à une suite d'observations naturalistes, mais que, comme les autres disciplines des sciences biologiques, elle s'appuie sur des concepts et des hypothèses scientifiques qu'il convient d'énoncer clairement et rigoureusement, et, le cas échéant, de tester par des approches expérimentales.

Délimitation du sujet et notions attendues. L'intitulé du sujet demandait qu'on explicite la diversité des relations interspécifiques. Ainsi, il était nécessaire de proposer une méthode de classification rigoureuse. Le plus judicieux était de classer les relations interspécifiques en fonction du résultat de la relation sur chacun des partenaires en terme de valeur sélective : les relations à bénéfice réciproque (mutualisme), les relations où l'un des partenaires exploite l'autre (prédation au sens large), les relations négatives pour les deux partenaires (compétition) et les relations unilatérales (commensalisme et amensalisme) ou neutres. En complément, il était possible d'analyser la spécificité des relations ainsi que leur degré d'intimité. De nombreux candidats ont omis de traiter de la compétition, probablement du fait d'une lecture imprécise du sujet qui ne limitait pas les relations interspécifiques aux relations trophiques.

Ensuite, on attendait une analyse de la façon dont les relations interspécifiques participent à la dynamique des populations et des communautés. L'exploitation du modèle de Lotka-Volterra permettait de montrer que les relations interspécifiques sont modélisables et que leurs effets sur les paramètres démographiques peuvent être prédits. La dynamique des populations a rarement été présentée précisément, même si le graphe des oscillations des effectifs de populations liées dans une relation proie-prédateur ou proie-parasite était connu. On pouvait ici discuter de l'importance de la prédation clef de voûte qui assure indirectement la stabilité du réseau d'interactions. Les notions d'espèce clef de voûte, ingénieur et architecte ont souvent été présentées de manière confuse, les candidats confondant souvent les espèces clefs de voûte et architectes. On devait également introduire le lien entre relations interspécifiques et stabilité de l'écosystème, en démontrant par exemple que la complexité des réseaux de relations interspécifiques est positivement corrélée à la résilience de l'écosystème.

L'importance des relations interspécifiques dans la structuration des communautés était également un aspect central du sujet. Il était nécessaire ici de développer la notion de niche écologique potentielle et réalisée, en lien avec l'existence d'une ou plusieurs relations interspécifiques. Cela devait amener à décrire les processus d'exclusion et de répartition spatiale et temporelle des populations qui en découlent. C'était également le bon cadre pour

développer l'exemple de l'effet Janzen-Connell qui montre comment les réseaux de relations interspécifiques peuvent modifier la répartition spatiale des populations.

Enfin, on attendait une discussion autour de l'impact des relations interspécifiques sur le fonctionnement de l'écosystème, en particulier sur la circulation de matière et d'énergie. Ainsi, la circulation de matière au sein de l'écosystème repose en grande partie sur les interactions de prédation qui définissent chaînes et réseaux trophiques. Les chaînes et réseaux trophiques ont très souvent été décrits, mais le lien avec les interactions interspécifiques a souvent été abordé de façon trop succincte. Des concepts complexes, comme l'impact de la syntrophie sur les rendements trophiques, étaient toutefois présents dans un certain nombre de copies.

Le sujet a été mal délimité par un certain nombre de candidats qui ont donc proposé une à deux grandes parties hors sujet, une conduite particulièrement pénalisante. Par exemple, certains candidats ont mis l'accent sur les conséquences évolutives des relations interspécifiques en insistant sur la coévolution et ses mécanismes. D'autres candidats n'ont pas non plus su placer la limite à partir de laquelle l'espèce interagit avec d'autres espèces (directement ou indirectement), ou avec l'environnement (les relations interspécifiques étant alors très indirectes). Un nombre notable de candidats a par exemple disserté trop longuement sur les décomposeurs et leur rôle dans l'écosystème. Le cycle du carbone a été également plusieurs fois décrit en détail. A l'inverse, les candidats proposant trois grandes parties regroupant chacune une catégorie de relations interspécifiques ont éludé la dimension écologique du sujet. Chaque terme du sujet est important, autant que les liens logiques, et doit être pris en considération pour construire et structurer un exposé répondant à l'intitulé.

Introduction. L'introduction nécessite une définition précise des termes du sujet. Ainsi, une introduction « minimale » devait comporter une définition de la notion d'espèce et de celle d'écosystème. Peu de candidats ont présenté la notion d'espèce, alors même que certaines relations intimes peuvent conduire à des situations où les frontières entre espèces peuvent être discutées. L'introduction a quelquefois révélé une confusion entre biotope et biocénose. Il convient d'être rigoureux dans les définitions pour pouvoir aborder sereinement le sujet. Une absence de définitions était souvent le symptôme d'une problématique peu travaillée, ce qui conduisait à un développement peu organisé, avec une absence de transition entre parties et la juxtaposition d'idées sans lien les unes avec les autres. S'il était connu d'un certain nombre de candidats que la présence ou l'absence de ruminants par exemple provoquait un changement important de l'écosystème prairial, les copies partant d'une observation naturaliste ou expérimentale pour amener le sujet étaient trop rares. Enfin, il est surprenant que peu de candidats aient précisé que les exemples qu'ils allaient utiliser seraient préférentiellement issus de la prairie pâturée, que ce soit en introduction ou dans le reste de l'écrit.

Exemples et illustrations. Le jury tient à saluer l'association, dans une grande majorité de copies, d'une illustration ou d'un exemple à chaque argument de l'exposé. Nous rappelons qu'un même schéma ou exemple peut être exploité à divers moments de l'exposé. Ainsi, il pouvait être judicieux d'utiliser le même exemple pour illustrer la notion de réseau trophique,

puis les régulations de l'écosystème par les prédateurs. Cette stratégie pouvait éviter les redondances et permettre de dégager de façon plus convaincante les liens entre les différentes parties de l'exposé. En effet, il est souvent inutile de multiplier les exemples autour de la même idée. Ainsi, certaines copies ont usé de quatre exemples différents sur la seule symbiose. Il était difficile voir impossible de recourir à un tel niveau de détail sans se répéter et sans occulter d'autres aspects du sujet. A l'inverse, de nombreux candidats ont proposé un schéma bilan qui synthétisait parfois avec réussite les idées développées dans leur copie. Nous rappelons que toute illustration doit être soignée, comporter un titre et être légendée et explicitée : une illustration non compréhensible par les correcteurs ne sera pas valorisée. En outre, les exemples doivent être précis et décrits avec un vocabulaire scientifique approprié. Par exemple, la chaîne trophique Poacées → Vache → Homme a parfois été simplement citée, alors qu'il était nécessaire de préciser que ces trois éléments étaient respectivement les producteurs primaires, le consommateur primaire et le consommateur secondaire, et que les flèches représentaient des relations de prédation de type phytophagie et prédation au sens strict, respectivement, et donc la circulation de matière et d'énergie entre ces trois niveaux trophiques.

Les quelques copies dans lesquelles les exemples étaient absents se sont trouvées fortement pénalisées. Par ailleurs, certains exemples ne concernaient pas la prairie pâturée, alors que cette limitation était clairement exigée dans l'énoncé. Cela n'a toutefois pas été pénalisé s'il s'agissait d'une exception dans la copie et qu'aucun exemple issu de la prairie pâturée n'était aussi pertinent que celui choisi par le candidat.

Conclusion. Comme la plupart des sujets d'écologie, l'élaboration de cette synthèse imposait de choisir des limites et de les justifier : il fallait ainsi entre autres définir des limites aux notions d'espèce et de relation interspécifique. Ces limites devaient être discutées en conclusion, à la lumière des éléments apportés par l'exposé. On pouvait également discuter du positionnement de l'homme dans le cadre du sujet, ce que bon nombre de candidats ont su faire.

Sujet d'analyse de documents - Le système CRISPR/Cas : mécanisme, aspects évolutifs et applications

Ce sujet était conçu pour évaluer les compétences qui sont indispensables à l'activité de recherche et d'enseignement en Biologie. En particulier, le jury a valorisé les copies démontrant les compétences suivantes :

- **compétences d'analyse** : être capable d'observer et d'exploiter des résultats expérimentaux pour recueillir des informations ;
- **compétences d'analyse critique** : être capable d'identifier d'éventuelles sources de biais méthodologiques dans les expériences analysées, commenter les incertitudes statistiques sur les valeurs expérimentales quantifiées ;

- **compétences intégratives** : être capable de hiérarchiser de multiples informations et de les articuler entre elles pour dégager un modèle complexe, et confronter ce modèle à un éventuel modèle connu auquel il se rapporte ;
- **compétence de communication scientifique** : être capable de construire une argumentation rigoureuse, en maîtrisant les relations de causalité et en suivant une progression logique ; savoir appuyer son propos sur des représentations schématiques appropriées.

Ce sujet a permis une sélection efficace des candidats. En effet, si les parties I et II ont été globalement traitées par une large part des candidats, la partie III, qui abordait les aspects évolutifs du mécanisme CRISPR/Cas vus sous l'angle de la génétique des populations, n'a été traitée correctement que par quelques candidats. La partie IV a quant à elle été seulement survolée dans la plupart des copies.

Le jury souhaite donner quelques conseils généraux aux futurs candidats pour les sujets d'analyse de documents :

- **Lisez bien l'énoncé.** Cette remarque semble triviale, pourtant une lecture trop rapide des légendes ou de certains panels a souvent été à l'origine d'une interprétation partielle voire erronée des résultats. Par exemple, le rôle du système CRISPR/Cas comme mécanisme de résistance des procaryotes aux éléments génétiques étrangers était une donnée de l'énoncé, ce qui devait guider l'analyse des résultats expérimentaux. De même, les légendes des figures et les panels présentant les principes expérimentaux comportent des éléments indispensables à la bonne compréhension des expériences. Par exemple, dans la table 2, la propriété d'auto-épissage de l'intron était essentielle à la compréhension des résultats. Dans la figure 4, le fait que les sondes utilisées pour la révélation du Southern blot étaient de séquence complémentaire au génome du phage ainsi que la taille des fragments indiqués dans le panel A devaient permettre aux candidats de comprendre qu'on analysait le devenir du génome du phage au cours de l'infection, et non le génome bactérien.
- **Utilisez un vocabulaire précis et adapté** lorsque vous analysez un document. Dire qu'un facteur est « impliqué dans » ou « intervient dans » ou « a un rôle dans » n'a pas la même signification que « est nécessaire à », « est suffisant à » ou « est nécessaire et suffisant à ». Nous rappelons que le jury accorde une grande importance à la rigueur scientifique et donc au choix des termes utilisés. Lorsque l'exercice introduit une terminologie, les candidats sont invités à l'utiliser à la place de termes plus génériques. Par exemple ici, il est préférable d'utiliser le terme « ARNcr » introduit par l'énoncé plutôt que simplement « ARN », pour faire référence à cette espèce particulière d'ARN. Introduire des erreurs dans ces termes précis sera perçu comme un manque de rigueur par le jury (par exemple, « séquence espacée et répéteur » au lieu de « séquence espaceur et répétée »).
- **Nous vous encourageons à synthétiser les données des exercices sous forme de schémas.** Cependant, tout schéma doit être légendé et explicité : un schéma non compréhensible par les correcteurs ne sera pas valorisé.

- **Faites l'effort de dégager le but de chaque expérience, puis d'interpréter et de conclure après chaque expérience.** Si vous restez au simple niveau de l'analyse, vous aurez des difficultés à progresser dans la compréhension du phénomène biologique étudié. Inversement, une conclusion juste mais qui ne s'appuierait pas sur une analyse rigoureuse des résultats ne sera pas pleinement valorisée. Pour le commentaire des résultats expérimentaux, nous recommandons aux candidats de suivre le schéma classique suivant : (1) brève description de l'expérience et explicitation de son (ou de ses) but(s), (2) analyse factuelle des résultats, (3) interprétation en lien avec les buts énoncés, (4) conclusions (notamment s'il y a plusieurs figures) et éventuellement (5) énoncé d'hypothèses explicatives.
- **Indiquez clairement si les conclusions que vous donnez sont directement issues de l'analyse des documents et donc démontrées par les données expérimentales, ou s'il s'agit d'hypothèses spéculatives.** Si le jury encourage les candidats à proposer des hypothèses après l'interprétation des résultats, celles-ci ne doivent pas être reprises dans les questions suivantes comme des éléments démontrés, au risque d'aiguiller l'analyse dans une mauvaise direction.
- **Faites l'effort d'intégrer l'ensemble des données apportées par les différentes figures dans un modèle général que vous devez construire tout au long de l'exercice.** Les sujets sont construits pour permettre aux candidats, en faisant la synthèse de leurs interprétations successives, de proposer une vision d'ensemble du mécanisme étudié. Sauf cas particulier, les figures ne sont pas indépendantes les unes des autres. Certaines questions, en général à la fin de chaque partie, incitent les candidats à faire cet effort de synthèse (qui ne doit pas se limiter à résumer la question précédente), mais on attend que les différentes réponses d'un candidat soient cohérentes entre elles et s'appuient sur ou soient discutées en fonction des réponses précédentes. Si des incohérences sont notées par le candidat, le jury apprécie que celles-ci soient commentées.
- **Maîtrisez les techniques de base de biologie,** tant dans leur protocole que dans leur principe. Dans ce sujet, de nombreuses confusions entre northern blot et Southern blot ont été faites. La notion de spécificité dans ces deux techniques (par rapport à une révélation à l'aide d'un intercalant des acides nucléiques comme le BET par exemple) est souvent mal comprise. Le **vocabulaire spécifique** des techniques est souvent aussi approximatif : par exemple, dans les électrophorèses, on parle de piste (et non pas de colonne) et de bandes (et non pas de taches, traits, ou autre). Les **unités de longueur ou de poids des molécules biologiques** doivent aussi être maîtrisées et utilisées de façon systématique : les Da (ou plus généralement kDa) pour les protéines, les pb (« paires de bases ») pour les acides nucléiques. Une manière de progresser consiste à s'exercer sur des sujets d'annales en s'appuyant sur la lecture des rapports de jury correspondants.

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

Cette première partie avait pour objectif la mise en évidence du mécanisme CRISPR/Cas et de ses propriétés essentielles, à savoir : (i) qu'il s'agit d'un mécanisme général de résistance vis-à-vis des transferts horizontaux de gènes chez les procaryotes, que ce soit la transduction (infection par un phage) ou la conjugaison ; (ii) que la résistance est dépendante de séquences espaceurs au sein du locus CRISPR identiques à des séquences présentes dans ces éléments génétiques étrangers ; (iii) que ces séquences espaceurs peuvent être acquises de façon dynamique ; et (iv) que la résistance résulte du clivage de l'ADN entrant au niveau de la séquence identique à la séquence espaceur du locus CRISPR.

Les questions 1 et 2 permettaient de corréliser la résistance à l'infection par un phage à l'identité de séquence entre une ou plusieurs séquences espaceurs du locus CRISPR et le génome du phage. La difficulté principale de ces questions résidait dans le fait que l'énoncé apportait de nombreuses informations que les candidats devaient trier et hiérarchiser pour en dégager les conclusions essentielles. Nous soulignons le fait qu'il est judicieux d'adopter une rédaction claire et synthétique et qu'il est inutile de paraphraser l'énoncé.

La figure 1 permettait de remarquer que les locus CRISPR des souches résistantes présentaient tous des séquences répétées et espaceurs supplémentaires par rapport à la souche WT (et non en moins comme certains candidats l'ont indiqué), situées entre la séquence L et la première séquence répétée de la souche WT. On pouvait noter que le nombre de séquences répétées et espaceurs supplémentaires variait d'une souche à l'autre. La figure 2 indiquait que les séquences espaceurs supplémentaires repérées en figure 1 étaient identiques à une séquence du phage utilisé pour l'infection, pointant vers une origine commune à ces séquences. L'acquisition de ces séquences par les bactéries ayant fait suite à l'infection par les phages, on pouvait faire l'hypothèse que ces séquences espaceurs dérivait de séquences de phages. Une analyse plus fine permettait de conclure qu'une similarité sans identité ne semblait pas suffisante pour conférer une résistance.

On pouvait également remarquer que lorsque plusieurs séquences espaceurs identiques aux séquences de phages avaient été acquises (cas des souches VIII et IX), la sensibilité relative était très basse, suggérant un effet additif des différentes séquences S.

La question 3 devait permettre de valider l'hypothèse proposée dans la question 2, à savoir que la résistance au phage était bien conférée par les séquences S acquises. Le jury a particulièrement valorisé les copies explicitant le fait que les séquences S supplémentaires étaient à la fois nécessaires et suffisantes pour conférer la résistance, ce qui pouvait être conclu de l'analyse des souches X et XI, respectivement. Ainsi, comme souvent, c'est la rigueur de l'analyse plus que la conclusion elle-même qui a été évaluée dans cette question.

La question 4 permettait d'étendre les résultats obtenus précédemment sur la résistance aux phages à d'autres types d'éléments génétiques étrangers, à savoir les plasmides acquis par conjugaison. Un certain nombre de candidats a mal compris cette expérience et a cherché à y démontrer la transmission du locus CRISPR par conjugaison par exemple, ou a pensé que l'on mesurait de nouveau la résistance aux phages. En fait, la comparaison de

l'efficacité d'acquisition du plasmide pNes(wt) par les souches *S. epidermidis* WT et *S. epidermidis* mut (délétion du locus CRISPR) permettait de conclure que le locus CRISPR était nécessaire pour résister à la conjugaison. Par analogie à l'infection par les phages, on pouvait faire l'hypothèse que l'espaceur identique au gène *nes* porté par le plasmide était responsable de la résistance à la conjugaison. Cette hypothèse était alors validée par la comparaison de l'efficacité d'acquisition des plasmides pNes(wt) et pNes(mut) par la souche *S. epidermidis* WT : l'identité de séquence entre le gène *nes* du plasmide et de l'espaceur était bien nécessaire à la résistance à la conjugaison. Les mutations introduites étant silencieuses, on en déduisait que ce n'était pas la séquence de la protéine codée mais bien du gène *nes* lui-même qui était importante.

Dans cette question encore, une construction rigoureuse de l'analyse était indispensable à la bonne interprétation des résultats. On attendait en effet des candidats qu'ils posent des hypothèses et qu'ils définissent une stratégie d'analyse permettant de les tester, et non qu'ils commentent de façon linéaire et systématique chaque colonne du tableau. Une telle rigueur méthodologique était nécessaire pour éviter les conclusions erronées ou non pertinentes.

Dans l'expérience de la **question 5**, on cherchait à définir si le mécanisme CRISPR/Cas ciblait directement l'ADN étranger ou indirectement l'ARN issu de la transcription de l'ADN étranger. En effet, l'intron auto-épissable était présent dans l'ADN plasmidique (donc la séquence de l'ADN du plasmide n'était pas identique à la séquence de l'espaceur bactérien) mais pas dans l'ARN (donc la séquence de l'ARN issu du plasmide était identique à la séquence de l'espaceur bactérien). Si on observait une résistance à la conjugaison par pNes(intron), c'était donc que le mécanisme ciblait l'ARN. En revanche, si on n'en observait pas, c'est que le mécanisme ciblait l'ADN uniquement. Les résultats montraient ainsi que le mécanisme ciblait directement l'ADN étranger. L'objectif de cette question a été rarement compris.

La première partie de la question 6 visait à mettre en évidence une corrélation entre la résistance au phage et le clivage rapide de l'ADN du phage au cours de l'infection, au niveau de la séquence identique à l'espaceur S7. Cette question nécessitait de bien comprendre la construction présentée en figure 4A, qui indiquait entre autre que les sondes 1 et 2 reconnaissaient le génome du phage (et non le génome bactérien comme certains candidats l'ont pensé). Elle appelait également une analyse patiente et complète des résultats de Southern blot.

La suite de la question 6 était une question de synthèse visant à résumer la partie I. Seuls de rares candidats ont compris que le mécanisme CRISPR/Cas était un mécanisme général de résistance vis-à-vis d'éléments génétiques étrangers, que ceux-ci soient acquis suite à l'infection par un phage ou par conjugaison. On attendait également que les candidats fassent apparaître le fait que le phénotype de résistance était dépendant de séquences espaceurs au sein du locus CRISPR identiques à des séquences présentes dans les éléments génétiques étrangers, et que ces séquences espaceurs pouvait être acquises de façon dynamique au cours de l'infection par un phage. Enfin, il fallait noter que la résistance reposait sur le clivage de l'ADN entrant au niveau de la séquence identique à la séquence

espaceur du locus CRISPR. Les candidats qui avaient dégagé peu de conclusions aux questions précédentes ont difficilement traité cette question.

Partie II - Modalités moléculaires du mécanisme CRISPR/Cas

Cette deuxième partie avait pour objectif de caractériser les bases moléculaires du mécanisme CRISPR/Cas, c'est-à-dire de comprendre comment le locus CRISPR/Cas permet la reconnaissance de l'ADN étranger puis son clivage.

La question 7 portait sur l'analyse d'un document simple présentant différentes souches bactériennes sensibles ou non à l'infection par le phage λ . Le principal écueil a résidé dans une analyse du document sans avoir clairement identifié la question à laquelle celui-ci cherche à répondre. Ainsi, la plupart des candidats ont correctement identifié les souches sensibles et résistantes, mais seuls quelques-uns ont correctement conclu que les gènes *casABCDE* et *cas3* étaient suffisants pour conférer la résistance.

La question 8 portait sur des expériences plus complexes et a très souvent donné lieu à des conclusions superficielles. Le document permettait de conclure que le locus CRISPR était à l'origine de plusieurs espèces d'ARNcr produits à partir d'un transcrit initial, la région centrale de chaque espèce d'ARNcr portant la séquence d'un espaceur donné. L'obtention de ces ARNcr à partir du transcrit initial par clivage ou épissage devait être proposée en hypothèse et étayée par les observations du document 6.

La question 9 devait permettre aux candidats de faire appel à leurs connaissances pour proposer une comparaison entre les ARNcr et les ARN interférents des eucaryotes. La plupart des copies ont répondu à cette question, mais les justifications complètes étaient rares. En effet, on attendait de lire que les ARNcr étaient (i) de petits ARN (ii) maturés à partir de transcrits plus longs et (iii) de séquence complémentaire à certaines séquences présentes dans les éléments génétiques auxquels ils confèrent une résistance.

La question 10 portait encore une fois sur l'analyse simple de plusieurs souches bactériennes dont la capacité à produire les ARNcr variait. On pouvait ainsi identifier le gène *casE* comme un facteur nécessaire à la production des ARNcr, et les gènes *casB*, *casC* et *casD* comme des facteurs y contribuant mais non indispensables. Ici aussi, il était nécessaire de formuler correctement la question posée par l'expérience pour pouvoir interpréter correctement les résultats, au-delà de la simple mention de l'abondance variable des ARN dans chacune des pistes.

La question 11 visait à mettre en évidence la formation d'un complexe entre Cascade, un ARN (vraisemblablement un ARNcr) et l'ADN cible. La figure 8A a été généralement bien comprise, même si l'hypothèse que l'ARN correspondait à un ARNcr a été rarement formulée. Le principe de l'expérience de retard sur gel (figure 8B) a été bien compris en général, même si les justifications pourraient être plus rigoureuses.

En question 12, on cherchait à montrer le rapprochement physique entre Cas3 et CasA au cours de l'infection. Si le principe de l'expérience de complémentarité bimoléculaire a été bien compris dans l'ensemble, la conclusion biologique n'a été que rarement formulée.

L'ensemble des résultats des parties I et II devaient être synthétisés dans **la question 13**. Ce type de question de synthèse, qui doit permettre aux candidats de démontrer leur compréhension globale du processus étudié, est souvent traité de façon trop rapide et très imprécise. Il ne suffit pas de replacer de façon approximative les différents éléments vus auparavant : il faut réfléchir précisément aux liens logiques les unissant. Le jury attend une réponse complète dans laquelle les différents éléments sont judicieusement articulés.

Partie III - Aspects évolutifs du mécanisme CRISPR/Cas

Pour cette troisième partie, on quittait le domaine de la biologie moléculaire pour faire un détour par la biologie évolutive et la génétique des populations. Ces aspects du sujet ont été peu abordés par les candidats.

La question 14 demandait aux candidats de dégager les avantages et désavantages évolutifs du système CRISPR/Cas. Si l'avantage conféré par l'acquisition d'une résistance aux phages a été très souvent bien compris, le coût évolutif d'un tel système n'a été mentionné que par la moitié des candidats ayant traité la question. On pouvait en effet postuler que la résistance aux éléments génétiques étrangers constituait un frein aux brassages génétiques chez les procaryotes, brassages qui contribuent à l'adaptabilité de ces organismes à un environnement changeant.

La question 15 n'a que rarement été abordée, alors que les hypothèses acceptées par le jury étaient nombreuses et ne nécessitaient aucune connaissance particulière. On pouvait ainsi proposer que le système ait été inactivé par mutation du locus CRISPR lui-même (délétion ou mutation de la séquence espaceur *nes*), par mutation perte de fonction des gènes *cas* associés, ou encore par mutation de la séquence cible (séquence *nes*). Toute hypothèse pertinente a été valorisée. Le jury incite donc les candidats à ne pas hésiter à formuler des propositions, si tant est qu'elles soient biologiquement plausibles et en accord avec les données de l'exercice.

La question 16 permettait de montrer que le locus CRISPR des souches transconjugantes avait subi une ou plusieurs délétion(s) par rapport à la souche WT. Les réponses étaient

souvent trop peu précises avec la simple mention que le locus était « différent » chez ces souches. Autant que faire se peut, les candidats doivent fournir des quantifications des résultats analysés. Ici par exemple, on pouvait établir que les délétions avaient affecté 50, 100 et 200 pb chez les souches 1, 2 et 3, respectivement. Le lien avec le phénotype transconjugant des souches (c'est-à-dire le fait qu'elles ne résistent plus à la conjugaison) devait également être établi : on pouvait en effet postuler que la séquence espaceur *nes* au moins avait été délétée.

La question 17 portait sur une expérience de fluctuation sous la forme d'une application simple du cours de BCPST. Cette question ne présentait donc aucune difficulté particulière, mais pour autant, elle n'a pas été traitée de façon satisfaisante sur l'ensemble des copies. On attendait une brève justification de la réponse montrant que les candidats avaient compris le phénomène : les copies concluant simplement à la préexistence du génotype muté sans autres explications n'ont donc pas été valorisées.

La question 18 devait permettre d'évaluer la capacité des candidats à appliquer leurs connaissances en génétique des populations à cet exercice. Ici encore, les attendus étaient très proches du cours de BCPST, mais exigeaient une bonne compréhension de cette partie du programme et une certaine habileté en calcul formel. La principale source d'erreur a résidé dans l'oubli de l'étape de normalisation de la fréquence à l'issue de la génération $t-1$, normalisation nécessaire pour que la somme des fréquences des deux génotypes à la génération t soit égale à 1.

La question 19 a été très peu traitée également (environ 1 copie sur 10). L'importance de l'expérience témoin n'a été que très rarement comprise. En effet, dans l'expérience présentée à la figure 11, on cherchait à estimer la valeur sélective associée au génotype muté par des essais de compétition entre la souche bactérienne sauvage et la souche mutée transconjugante. Or ces deux souches différaient par deux paramètres : leur génotype d'une part, et le fait que l'une ait acquis un plasmide et l'autre non, d'autre part. Afin de déterminer uniquement l'effet du génotype, on avait donc réalisé l'expérience témoin qui permettait de tester la valeur sélective associée à l'acquisition du plasmide pNes uniquement (les deux souches mises en compétition ayant le même génotype). L'expérience témoin permettait également d'estimer le poids de la dérive génétique dans les résultats, chaque passage successif pouvant causer des fixations d'allèles aléatoires.

L'analyse de la figure 11, qui devait permettre de conclure que la perte du système CRISPR/Cas pouvait être délétère ou bénéfique selon les expériences, a rarement été menée à son terme. La discussion attendue en c) était souvent décevante. On pouvait par exemple spéculer que le locus CRISPR/cas pouvait évoluer de façon extrêmement dynamique au sein des populations bactériennes, avec l'acquisition de nouvelles séquences espaceurs lorsqu'une pression de sélection de type phage était exercée, une délétion du locus lorsque les pressions de sélection étaient en faveur de l'acquisition d'éléments génétiques étrangers (dans un environnement changeant en particulier), puis une nouvelle acquisition du locus par transfert horizontal.

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

La dernière partie de l'exercice abordait l'exploitation biotechnologique du système CRISPR/Cas en tant que stratégie innovante pour l'édition de gènes. Malgré la couverture médiatique importante de cette technologie au cours des derniers mois, seul un nombre très limité de candidats a correctement traité cette partie en faisant le lien entre le mécanisme biologique mis en évidence dans les premières parties de l'exercice et son application technologique. Nous invitons les candidats à entretenir leur culture scientifique en lisant de façon régulière la presse généraliste et en s'exerçant à faire le lien entre l'actualité scientifique et le contenu des programmes de BCPST, non seulement pour la préparation du concours mais également pour la préparation de leur formation à la recherche.

La question 20 demandait une analyse des transgènes utilisés dans l'expérience d'édition de gènes *in vivo*. Peu de candidats ont compris qu'en cellule eucaryote, il était nécessaire que SpCas9 soit adressée au noyau cellulaire (qui contient le génome à éditer), d'où l'utilisation d'une séquence de localisation nucléaire. Certains ont proposé que cette séquence permettait l'adressage du transgène, et non de la protéine codée par celui-ci, au noyau, ce qui était incorrect. Seules de rares copies ont correctement expliqué qu'en visualisant les cellules dans lesquelles le gène codant la GFP s'exprimait, on pouvait identifier les cellules dans lesquelles le vecteur ARNcr était entré efficacement, donc dans lesquelles l'édition du gène cible avait pu être réalisée. Enfin, si la plupart des copies ayant traité cette question ont judicieusement proposé d'utiliser un promoteur actif spécifiquement dans les neurones, rares sont celles ayant pensé à utiliser le promoteur de *Mecp2* lui-même pour restreindre encore l'expression de SpCas9 et de la GFP aux cellules exprimant *Mecp2*.

La question 21 a le plus souvent été traitée de façon superficielle pour une telle question de synthèse, probablement du fait de sa position en fin de sujet. De la figure 13, on pouvait tirer que les séquences issues des cellules GFP-positives (donc dans lesquelles l'édition avait pu avoir lieu) présentaient des mutations par rapport à la séquence sauvage (délétions, insertions ou combinaison des deux). Dans la majorité des cellules, un nombre de nucléotides non multiple de trois était affecté, aboutissant à un décalage du cadre de lecture et possiblement à l'apparition d'un codon stop prématuré. Puisque les parties I et II avaient démontré un clivage de l'ADN cible, on pouvait proposer que chez les eucaryotes, le clivage de l'ADN cible aboutissait à l'apparition de délétions et insertions lors de la réparation des cassures de l'ADN.

La figure 14 complétait ces résultats par la mise en évidence en immunofluorescence et western blot de l'inactivation partielle de l'expression de MeCP2 chez les souris ayant reçu

les vecteurs de transgénèse. Ainsi, on illustre le fait que la technologie CRISPR/SpCas9 permettait effectivement l'édition de gène *in vivo*.

S'appuyant sur ces observations, on attendait alors en **question 22** que les candidats citent un avantage de cette stratégie par rapport à la stratégie de *knock-out* classique. L'avantage principal de la stratégie CRISPR/SpCas9 réside dans le fait que les souris modifiées sont obtenues en quelques semaines seulement, contre plusieurs mois pour les souris *knock-out* classiques. Cependant, toutes les idées intéressantes ont été valorisées.

La **question 23** demandait un schéma bilan de l'exercice tout entier. Cette question a été très rarement traitée et de façon souvent approximative et imprécise lorsqu'elle l'était, du fait d'un manque de profondeur dans l'analyse des documents en amont. Très peu de candidats ont fait l'effort de reprendre l'intégralité des résultats de l'exercice. Les copies présentant un bilan global ont été valorisées. Dans ce type de question de synthèse, le jury invite les candidats à mettre l'accent sur les liens entre les éléments tirés des différentes parties de l'exercice. Ici par exemple, on pouvait mettre en lien le phénotype de résistance aux éléments génétiques étrangers mis en évidence dans la partie I et l'impact sur la valeur sélective mis en évidence dans la partie III, ou encore la reconnaissance puis le clivage de l'ADN cible mis en évidence dans les parties I et II et l'introduction de mutations dans le gène cible par la technologie SpCas9/ARNcr illustrée dans la partie IV.