

ÉPREUVE ÉCRITE DE BIOLOGIE

ENS : PARIS LYON CACHAN

Coefficients : **PARIS** : Option Biologie : 7 / Option Géologie : 4
LYON : Option Biologie : 8 / Option Sciences de la Terre : 4
CACHAN : 8

MEMBRES DU JURY : D. BUSTI, N. CAUDRON, A. CORBIN, M.-H. KRYSZKE, S. LE CROM, G. PEYROCHE, P. PLA, B. SCHNEIDER.

L'épreuve de biologie de cette année, d'une durée totale de 6 h, comportait un sujet de synthèse d'une durée conseillée de 2 h ainsi que deux sujets avec documents d'une durée conseillée de 1 h (sujet A) et 3 h (sujet B), construits autour d'une thématique commune : les ARN.

Une bonne gestion du temps était déterminante puisque les points attribués à chaque sujet étaient proportionnels aux durées conseillées. Il fallait par conséquent éviter de consacrer trop de temps au sujet de synthèse, de façon à ne pas passer à côté de résultats importants lors de l'analyse des documents. A ce propos, nous tenons à saluer la performance des quelques candidats qui ont réussi à aborder et à comprendre la totalité de l'épreuve malgré la difficulté des dernières questions.

Les sujets avec documents, de difficulté croissante, ont permis une bonne sélection des candidats :

- le sujet A ('le RNA silencing chez les plantes') a souvent été abordé dans sa quasi-totalité.
- le sujet B-1 ('Contrôle de qualité des ARN au moment de leur exportation vers le cytoplasme') était d'une difficulté moyenne. Il comportait notamment une question assez sélective (question 16) sur laquelle une bonne moitié des candidats a buté.
- le sujet B-2 ('Contrôle de qualité des ARN au moment de leur traduction') était plus difficile, aussi bien dans l'analyse des documents que dans la construction du modèle final. Beaucoup ont décroché aux questions 26-27 qui se sont révélées très discriminantes.

Comme chaque année, nous souhaitons donner quelques conseils pour les sujets d'analyse :

- Lisez bien l'énoncé de façon à ne pas laisser de côté des informations nécessaires à l'interprétation des résultats.
- Répondez précisément aux questions : le plus souvent, les questions guident le raisonnement et attirent l'attention sur certains points importants qui pourraient être négligés par une lecture trop rapide des résultats.
- Analysez et interprétez les résultats expérimentaux, mais ne vous attachez pas à simplement les décrire, sauf indication contraire de l'énoncé. Évitez de « paraphraser » les graphiques ou les résultats d'électrophorèse : « dans la piste 1, on voit une tache noire... ».
- Séparez clairement les conclusions validées par les expériences d'éventuelles hypothèses qu'elles peuvent susciter, mais qui ne sont pas formellement démontrées.
- Faites la différence entre des résultats significatifs et des résultats qui ne le sont pas.
- Ne vous laissez pas influencer par vos connaissances propres, afin de ne pas déduire plus de choses que les résultats expérimentaux ne le permettent.
- Soyez rigoureux dans votre analyse en faisant référence aux expériences de contrôle et en donnant des valeurs chiffrées aux différences observées.
- Soyez le plus concis possible dans la rédaction : il n'est pas nécessaire qu'une réponse soit très longue pour qu'elle soit correcte.
- Utilisez un vocabulaire scientifique précis. Par exemple, la phrase « le produit du gène A contrôle l'activité du gène B » est moins informative que la phrase « le produit du gène A active l'expression du gène B ».

Sujet de synthèse – Les ARN

Il était attendu dans ce sujet de réunir et d'organiser des notions disséminées à différents endroits du programme (structure des acides nucléiques, transcription, traduction, génomes viraux à ARN), et de les présenter de façon claire et concise, en adoptant une approche expérimentale.

Nous tenons à saluer les efforts des candidats sur la présentation qui a été globalement soignée, même si certaines copies manquaient d'illustrations ou comportaient des schémas quelque peu fantaisistes (schéma d'une aminoacyl-ARNt synthétase par exemple). La rédaction a été beaucoup plus fréquemment négligée. Les fautes d'orthographe trop nombreuses, les erreurs grossières de français et un style trop finaliste (« Les ARN agissent dans le but de... ») ont été pénalisés. Attention au vocabulaire scientifique impropre (« l'ADN, une molécule séquencée ») et aux néologismes (« inxon », « transcriptique ») !

Dans la rédaction de l'introduction, la problématique a rarement été bien amenée : pour y parvenir, il était possible d'adopter une approche historique. Nous avons par ailleurs remarqué que beaucoup de conclusions ressemblaient à des introductions. Nous rappelons qu'il est essentiel en conclusion de faire ressortir les points forts du devoir (en évitant les banalités), puis de proposer une ouverture (par exemple : utilisation des ARNr pour la construction de l'arbre phylogénétique du vivant, utilisation d'ARN antisens pour inactiver un gène et comprendre son rôle, construction de banques d'ADN complémentaire à partir des ARNm cellulaires).

De manière générale, la démarche ne s'est pas suffisamment appuyée sur des résultats expérimentaux. Bien que les expériences d'hybridation ADN/ARNm aient été mentionnées dans plusieurs copies, rares sont celles qui ont décrit le système expérimental (cellules d'oviducte de poule synthétisant l'ovalbumine). Le plan choisi n'a pas toujours été logique : par exemple, une présentation de la structure des ARN s'imposait avant d'aborder les mécanismes de leur synthèse.

Si certaines notions ont presque toujours été abordées (structure des ARN, transcription, traduction), d'autres ont souvent été oubliées ou mal traitées (dégradation/stabilité comparée des ARNm eucaryotes/procaryotes, caractéristiques du code génétique, ARN viraux comme supports de l'information génétique, flottement dans la reconnaissance codon/anticodon...). Quelques hors-sujets ont été commis (structure de l'ADN, notion de virus...). Enfin, si beaucoup de candidats connaissent la structure des ARN, plusieurs erreurs ont été commises :

- Exons et introns ne se définissent pas comme étant respectivement codants et non codants, mais comme les séquences que l'on retrouve ou non dans l'ARNm mature. Un intron est obligatoirement situé entre deux exons.
- Trop de candidats mentionnent l'épissage alternatif, sans savoir de quoi il s'agit (principe, spécificité tissulaire). Certains même pensent que l'épissage alternatif est la règle ! Attention, en aucun cas l'épissage n'aboutit à modifier l'ordre des exons dans l'ARNm.
- Certains confondent épissage (des exons) et excision (des introns).
- De nombreux candidats mélangent les mécanismes d'initiation et de terminaison de la transcription chez les eucaryotes et les procaryotes. Très peu définissent clairement les sites importants des promoteurs procaryotes (site -35 où se fixe le facteur σ , site -10 où se fixe l'ARN polymérase et site d'initiation) et eucaryotes (promoteur minimal : site d'initiation et boîte TATA où se fixent les facteurs généraux de la transcription et l'ARN polymérase II, et séquences régulatrices distales : *enhancers*, *silencers*, éléments de réponse où se fixent les facteurs régulateurs).
- Certains se trompent dans le sens de synthèse 5'→3' des ARN.
- Des candidats ne connaissent pas le fonctionnement de l'opéron lactose et ont confondu promoteur et opérateur.
- Certains décrivent le processus de répllication de l'ADN en prétendant traiter de sa transcription.

Sujet A – Le ‘RNA silencing’ chez les plantes

Ce sujet avait pour objectif de mettre en évidence l’effet inhibiteur de l’ARN *MIR172* (un ARN non codant de 21 nucléotides) sur la traduction de l’ARNm *AP2* et de comprendre son rôle dans le contrôle de l’identité des pièces florales chez *Arabidopsis thaliana*. Pour cela, il fallait analyser les phénotypes de fleurs mutantes, les niveaux d’expression de gènes visualisés par électrophorèse et des profils d’expression de gènes obtenus par hybridation *in situ*.

Les cinq premières questions permettaient de déduire la boucle de régulation reliant les gènes *MIR172*, *AP2* et *AG*. A la question 1, la majeure partie des candidats a su déduire la classe des gènes *AP2* et *AG* (classes A et C, respectivement). Cependant, beaucoup d’entre eux se sont basés sur une analyse sommaire des phénotypes des fleurs mutantes et n’ont pas remarqué le remplacement des pétales et des sépales par des carpelles surnuméraires chez les fleurs *apetala2*, ainsi que le remplacement des étamines et des carpelles par des sépales et des pétales chez les fleurs *agamous*. Les données supplémentaires sur les domaines d’expression des gènes *AP2* et *AG* chez les plantes sauvages et mutantes, permettaient de conclure que les activités des gènes *AP2* et *AG* s’excluent mutuellement par inhibition réciproque (question 2). L’analyse des résultats des profils d’expression des gènes *MIR172* et *AP2* chez les plantes P-*MIR172* et les plantes mutantes *agamous* a été abordée par beaucoup, mais de manière assez disparate. Si de nombreux candidats ont vu l’effet inhibiteur de l’ARN *MIR172* sur l’expression du gène *AP2* au niveau traductionnel, peu d’entre eux font référence aux témoins (ARN totaux, ARN *UBQ5* ou protéine PEPC), beaucoup ne voient pas l’effet activateur du gène *AG* sur le gène *MIR172* et certains interprètent mal le phénotype des plantes *agamous*. On déduisait de l’ensemble des données que le gène *AG* active le gène *MIR172*, lequel inhibe (au niveau traductionnel ou post-traductionnel) le gène *AP2*, lequel inhibe en retour le gène *AG*. Beaucoup de candidats ont replacé ces interactions dans les différentes zones de la fleur, alors qu’ils n’avaient, à ce stade de l’exercice, aucun argument expérimental sur le domaine d’expression du gène *MIR172*.

Les questions 6 et 7 avaient pour but de montrer que l’effet inhibiteur de *MIR172* sur l’expression du gène *AP2* passe par l’hybridation de l’ARN *MIR172* à l’ARNm *AP2*. Dans le système expérimental utilisé, les mutations introduites dans le gène *AP2* (gène *AP2m*) ont pour conséquence de bloquer toute hybridation entre l’ARN *MIR172* et l’ARNm *AP2*, sans toutefois changer la séquence d’acides aminés et donc la fonctionnalité de la protéine *AP2* (les mutations sont silencieuses). Les conclusions tirées de la figure 4 se sont rarement appuyées sur une analyse rigoureuse et chiffrée des résultats expérimentaux. Certains candidats ont répondu à partir de leurs connaissances du mécanisme d’action des ARN interférents, sans analyser les résultats. La principale difficulté était de quantifier la baisse de l’effet inhibiteur de l’ARN *MIR172* sur la traduction de l’ARNm *AP2m* (si l’on tient compte du niveau d’expression des ARNm *AP2* et *AP2m*, l’effet inhibiteur mesuré chez les plantes P-*AP2m* est deux fois moindre que chez les plantes P-*AP2*).

Enfin, les questions 8 et 9 permettaient de comprendre comment l’évolution du profil d’expression du gène *MIR172* au cours du développement de la fleur sauvage d’*Arabidopsis thaliana* (expression dans tout le méristème floral au stade 1, puis seulement dans les ébauches des pièces reproductives au stade 7) est susceptible de contrôler le patron d’expression des gènes *AP2* et *AG*, et par là-même l’identité des pièces florales. Les candidats ont relativement bien traité ces questions, mais ont bien souvent tiré des conclusions sur l’origine du patron d’expression de l’ARN *MIR172* dans le méristème floral, alors que les résultats expérimentaux ne permettaient pas de le faire.

Sujet B – Contrôle de la qualité des ARN chez les Vertébrés

Ce sujet, assez long, avait pour but d'aborder quelques aspects du contrôle de qualité des ARN au moment de leur exportation vers le cytoplasme (partie B-1) et de leur traduction (partie B-2). Les parties B-1 et B-2 n'étaient pas totalement indépendantes dans la mesure où l'interprétation des derniers résultats de la partie B-2 (section B-2.2) devait s'appuyer sur les conclusions obtenues à la partie B-1. Toutefois, l'énoncé avait été construit de telle sorte que l'essentiel de la partie B-2 (section B-2.1) puisse être traitée indépendamment de la partie B-1. A l'issue de chacune des parties, les candidats devaient schématiser la cascade d'événements assurant le contrôle de qualité des ARN. Si certains sont parvenus à proposer un modèle satisfaisant à la question 22, très rares sont ceux qui en proposent un à la question 38. Ces questions, assez difficiles, exigeaient d'intégrer l'ensemble des conclusions dans un modèle cohérent. Seuls les candidats les plus persévérants ont été fortement récompensés.

B-1. Contrôle de la qualité des ARN au moment de leur exportation vers le cytoplasme

Les questions 10 à 13 permettaient de comprendre le principe du contrôle de qualité des ARN au moment de l'exportation. De manière surprenante, la question préliminaire (question 10) s'est révélée discriminante, puisque beaucoup ont répondu de manière incorrecte ou non exhaustive. Il était attendu du candidat qu'il mentionne les trois conséquences possibles de l'absence d'excision d'un intron : (1) l'intercalation de nouveaux acides aminés dans la protéine et/ou (2) l'apparition d'un codon stop prématuré donnant une protéine tronquée et/ou (3) le décalage de la phase de lecture de l'exon qui suit immédiatement l'intron à l'origine de substitutions d'acides aminés. La question 11 était une simple application du cours : on demandait simplement d'identifier l'ARNm *TPI* (bande 1) et l'intron excisé (bande 2) comme sous-produits de l'épissage de l'ARNpm *TPI*. Pour l'identification, les candidats pouvaient indifféremment utiliser les schémas de la figure 6.A représentés à l'échelle ou suivre la localisation subcellulaire des bandes 1 et 2. Quelques uns nomment la bande 1 de manière maladroite (Δ i-ARN *TPI* au lieu d'ARNm *TPI*) alors que d'autres identifient des sous-produits d'un épissage alternatif : exon 1 (bande 1) et exon 2 (bande 2). Cette dernière réponse n'a pas été acceptée dans la mesure où les deux exons d'un ARNpm possédant un seul intron sont obligatoirement réunis dans l'ARNm mature produit par épissage. A la question 12, l'analyse des résultats permettaient de conclure que les ARNm *TPI* issus de l'épissage (bande 1) sont exportés plus rapidement vers le cytoplasme que les ARNm *TPI* non issus de l'épissage (Δ i-ARN *TPI*). Le fait que le déroulement de l'épissage favorise l'exportation des ARN correctement épissés (ARNm) au détriment des ARNpm ou des ARNm partiellement épissés (susceptibles de donner des protéines non fonctionnelles au moment de la traduction) est à la base du contrôle de qualité des ARN au moment de l'exportation.

Les questions 14 à 20 mettaient en lumière les bases moléculaires de ce contrôle de qualité. Les expériences de co-immunoprécipitation des ARNm *TPI* issus ou non de l'épissage en présence de différents anticorps (question 14) montrent que la fixation des protéines d'exportation Y14, DEK et REF est dépendante de l'épissage et qu'elle s'effectue spécifiquement sur le fragment médian de l'ARNm *TPI*, contrairement à celle de la protéine hnRNP A1 qui se lie de manière non spécifique aux deux types d'ARN. Cette fixation différentielle peut ainsi rendre compte des différences de vitesse d'exportation observées pour ces deux types d'ARN dans la mesure où plus le nombre de protéines impliquées dans l'exportation est élevé, plus la vitesse d'exportation sera grande (question 15). Si presque tous les candidats ont abordé ces questions, peu ont su faire ressortir les résultats significatifs. Ceux qui ont mal lu l'énoncé ont conclu que ce sont les anticorps qui se fixent aux ARN (!), que les protéines Y14, REF et DEK se fixent au niveau de l'intron de l'ARN *TPI* épissés ou encore que la fixation de ces protéines freine l'exportation des ARNm *TPI* issus de l'épissage par encombrement stérique. Les questions suivantes, beaucoup plus discriminantes, permettaient de localiser plus précisément sur l'ARN le site de fixation de ces protéines

d'exportation. Les expériences de digestion à la RNase H en présence de différents oligonucléotides (questions 16 et 17) permettaient d'identifier un site d'interaction situé environ 24 nucléotides en amont de la jonction exon-exon, uniquement pour les ARNm *TPI* issus de l'épissage (la fixation d'une ou plusieurs protéines sur ce site protège l'ARNm de la digestion). Le même type d'expérience mené avec le gène *PIP85.B* (question 18) montre par ailleurs que l'insertion de 12 nucléotides en amont de la jonction exon-exon ne modifie pas le profil de digestion à la RNase H (le site protégé n'est pas décalé par rapport à la limite exon-exon), ce qui montre que cette fixation est indépendante du contexte nucléotidique. Les questions 19 et 20 montrent que les protéines Y14, DEK et REF (mais pas hnRNP A1) interagissent avec un site d'au plus 8 nucléotides s'étendant sur les positions -24 et -20 en amont de la jonction exon-exon.

Les expériences de co-immunoprécipitation des ARNm β nucléaire et cytoplasmique (question 21) permettaient de suivre *in vivo* l'évolution de la composition du complexe EJC au moment de l'exportation des ARNm. Les résultats montrent que, contrairement à Y14, les protéines DEK et REF n'interagissent plus avec les ARNm β cytoplasmiques issus de l'épissage : elles se dissocient donc de ces mêmes ARNm au moment de leur exportation. De nombreux candidats ont proposé de manière arbitraire un ordre de dissociation, alors qu'il était possible de le faire de manière argumentée : par exemple, DEK se dissocierait des ARN dans le noyau juste avant leur exportation dans la mesure où c'est une protéine exclusivement nucléaire.

Pour terminer, un schéma fonctionnel de l'évolution des interactions ARN/protéines depuis l'épissage jusqu'au transport des ARNm dans le cytoplasme était demandé (question 22). Certains candidats ont confondu le complexe d'épissage (*spliceosome*) qui agit au niveau des jonctions exon-exon, avec le complexe EJC qui se dépose au moment de l'épissage en amont de chacune de ces jonctions. L'évolution de la composition du complexe EJC a rarement été représentée correctement, la plupart des candidats proposant une dissociation de la protéine Y14 juste après l'exportation de l'ARNm.

B-2. Contrôle de la qualité des ARN au moment de leur traduction

La question 23 était déterminante pour la compréhension de la suite de l'exercice. Les résultats montrent que l'introduction artificielle d'un codon stop prématuré en position 189 dans la séquence codante du gène *TPI* (mutation non sens '189 Ter') induit une baisse drastique de la quantité d'ARNm *TPI* portant cette mutation. Cet effet est dû à la dégradation de l'ARNm *TPI* portant la mutation et non à un défaut de transcription du gène puisque le gène témoin de β -globine placé sous le contrôle du même promoteur et introduit dans les mêmes conditions s'exprime aussi fortement que le gène *TPI* sauvage. Il y a donc un contrôle de qualité des ARN au moment de la traduction qui prévient *in vivo* les effets délétères liés à l'apparition d'une mutation non sens risquant de donner une protéine tronquée non fonctionnelle. Les candidats qui ne se sont pas souciés du témoin ont souvent attribué cet effet à une transcription moins importante du gène. D'autres ont avancé sans argument valable que la mutation non sens induit un arrêt prématuré de la transcription, ce qui ne peut être le cas puisque l'ARNm *TPI* muté présente la même taille que l'ARNm sauvage.

Les questions 24 à 28 mettaient en lumière le fait que la machinerie cellulaire est capable de discriminer les codons stop prématurés (mutations non sens) des codons stops normaux terminaux, en fonction de leur distance à la dernière jonction exon-exon, et d'induire ou non la dégradation des ARNm. Les résultats obtenus à la question 24 suggèrent que (1) une limite existe entre les codons non sens qui induisent une dégradation de l'ARNm (mutations '189 Ter' et '192 Ter') et les codons non sens qui ne l'induisent pas (mutations '195 Ter', '198 Ter' et '208 Ter') – dans cette hypothèse, cette limite se situerait entre les codons 192 et 195 – ou que (2) l'effet différentiel observé pourrait dépendre du contexte nucléotidique dans lequel sont introduits les mutations non sens, indépendamment de leur position par rapport à une région déterminante (seconde hypothèse).

Les résultats obtenus avec les constructions '208 Ter + 39 pb' et '208 Ter + 57 pb' (question 25) infirment la seconde hypothèse puisque l'on obtient un effet différentiel alors que le contexte nucléotidique de la mutation reste le même, et démontrent que la position relative de la mutation '208 Ter' par rapport à une région située en aval du codon 208 est déterminante. Puisque la délétion des codons 211 à 230 n'a que peu d'effet sur le niveau d'expression des ARNm *TPI* portant la mutation '189 Ter' (question 26), cela montre que cette région n'est pas déterminante dans l'induction de la dégradation et que la région déterminante ne se situe pas en aval du codon 230. On pouvait alors proposer la règle selon laquelle (question 27) la limite entre les codons non sens qui induisent une dégradation de l'ARNm et les codons non sens qui ne l'induisent pas se situe à environ 50 nucléotides (entre 46 et 51 nucléotides d'après les résultats expérimentaux présentés ici) en amont de la dernière jonction exon-exon sur la séquence de l'ARNm. La question 28 permet de confirmer la règle précédente (le codon stop terminal des gènes possédant un voire deux exons non codants est situé dans presque tous les cas à une distance inférieure à 50 pb en amont de la dernière jonction exon-exon) et de proposer la machinerie de traduction (ribosome et facteurs associés) comme acteur susceptible de reconnaître une mutation non sens.

Les questions 29 à 34, rarement abordées, permettaient de reconstituer à partir d'un système artificiel astucieux, les étapes de la réponse induite par une mutation non sens. Les résultats obtenus avec les constructions β -6MS2 et β -TAC-6MS2 montrent que les protéines hUpf3a et hUpf3b sont chacune capable d'initier la dégradation des ARNm, mais à condition qu'elles soient recrutées sur l'ARNm en aval du codon stop (questions 29 et 30). Les différences observées peuvent s'expliquer par le fait que les ribosomes déplacent les protéines hUpf3 lors de la traduction des ARNm β -TAC-6MS2 mais pas des ARNm β -6MS2 (question 31). L'utilisation d'une version de hUpf1 inhibant l'activité du complexe hUpf3 permet de montrer que la dégradation des ARNm causée par l'apparition d'une mutation non sens, dépend en quasi-totalité de l'activité de ce complexe (question 32). Les expériences menées avec les gènes β et β -39, en présence ou en l'absence de cycloheximide, démontrent que la traduction est requise pour la dégradation des ARNm induite par une mutation non sens (question 33). Les résultats obtenus avec β -6MS2 précisent que l'activité du complexe hUpf3 est dépendante de la traduction. D'où le modèle suivant : lorsque la machinerie de traduction reconnaît un codon stop situé au moins 50 nucléotides en amont de la dernière limite exon-exon, celle-ci transduit un signal d'activation au complexe hUpf3, lequel induit la dégradation de l'ARNm portant la mutation (question 34).

Les résultats de la partie B-2.2 montrent que l'épissage du dernier intron est nécessaire à la dégradation des ARNm induite par l'apparition d'une mutation non sens (question 35) et que la protéine hUpf3b est associée *in vivo* aux protéines du complexe EJC (question 36). L'effet différentiel de la mutation '189 Ter' observé sur les ARNm issus ou non de l'épissage trouve son explication dans le fait que la protéine hUpf3b est recrutée au niveau des complexes EJC exclusivement sur les ARNm épissés (question 37).

Le modèle final devait faire transparaître le rôle du complexe EJC dans la discrimination entre un codon stop provenant d'une mutation non sens et un codon stop normal (question 38) :

- Lorsqu'un ARNm ne possède pas de codon stop situé à 50 nucléotides en amont de la dernière jonction exon-exon, la machinerie de traduction déplace tous les complexes EJC (y compris les protéines hUpf3a/b) et ne transduit aucun signal de dégradation en fin de traduction. Dans ce cas, le codon stop est interprété comme un signal de terminaison de la traduction.
- Lorsqu'un ARNm possède un codon stop se situant au moins 50 nucléotides en amont de la dernière jonction exon-exon, un ou plusieurs complexes EJC subsistent sur l'ARNm après le premier round de traduction. Le ribosome arrêté au niveau du codon stop anormal transduit alors un signal de dégradation au complexe hUpf3a/b situé en aval. Dans ce cas, le codon stop est interprété comme une mutation non sens.