

ÉCOLE NORMALE SUPÉRIEURE DE LYON

Concours d'admission session 2016
Filière universitaire : Second concours

COMPOSITION DE BIOLOGIE - BIOCHIMIE

Durée : 3 heures

L'usage de tout document et de calculatrice est interdit

Ce livret comprend 13 pages numérotées de 1 à 13

Cette épreuve comporte 2 parties qui peuvent être traitées de façon indépendante.

Les temps prévisionnels qu'il est suggéré de consacrer à chacune des parties, et qui serviront de base à l'élaboration du barème, sont indiqués ci-dessous :

•

Partie I 2h.

Mécanismes pathologiques liés aux mutations de *C9ORF72*

•

Partie II 1h.

Fonction de H2A.Z dans la mémoire

Des documents complémentaires sont présents en annexes.

Toute réponse devra être justifiée, même brièvement

Partie I.

Mécanismes pathologiques liés aux mutations de *C9ORF72*

Les dégénérescences lobaires fronto-temporales (DLFT) sont un ensemble de maladies caractérisées par la mort progressive des neurones des lobes frontaux du cerveau. Ces aires cérébrales sont associées entre autre à la personnalité, au langage, au comportement, à la gestion des émotions... Tous ces aspects peuvent être profondément altérés chez les patients, à des degrés divers. Certaines DLFT sont accompagnées de manifestations typiques d'une autre maladie dégénérative, la sclérose latérale amyotrophique (ALS, ou maladie de Charcot), caractérisée elle par des atteintes motrices touchant la musculature squelettique. On parle alors de maladie DLFT/ALS. Ces maladies peuvent toucher des sujets relativement jeunes, leur fréquence augmente avec l'âge. La maladie est aujourd'hui incurable et conduit à la mort en quelques années.

Bien que les mécanismes de pathogenèse soient encore mal compris, le gène *C9ORF72* est mutant dans près de 90% des cas, cas dans lesquels il est considéré comme la cause de la maladie DLFT/ALS. Vous venez de commencer un stage dans un laboratoire qui étudie les mécanismes pathologiques qui impliquent *C9ORF72*. La Figure 1 présente deux arbres généalogiques de deux familles ayant des membres souffrant de DLFT/ALS (adapté de (1) et (2)).

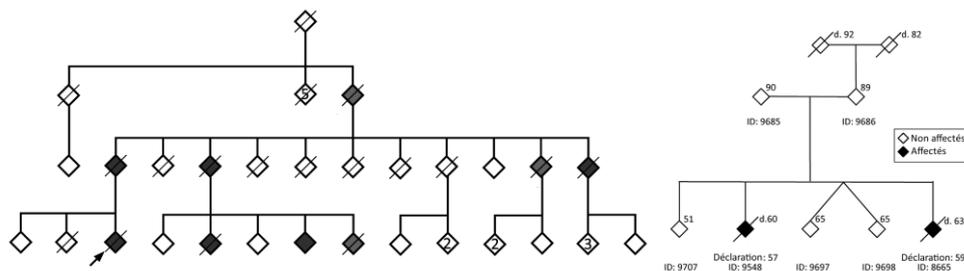


Figure 1. Arbres généalogiques de deux familles avec des membres atteints (losanges sombres) ou pas (losanges blancs) par DLFT/ALS. Le chiffre en haut d'un losange est l'âge de la personne, un losange barré indique une personne décédée, et la mention d. indique l'âge du décès le cas échéant. Déclaration : âge au diagnostic. Un chiffre X dans un losange indique qu'il existe X autres personnes dans la fratrie avec des caractères identiques. La flèche pointe vers le proband, c'est à dire la personne qui a motivé l'enquête génétique. Les conjoints n'apparaissant pas dans le diagramme sont sains.

Question 1. Quel mode de transmission de la maladie est suggéré par l'arbre de gauche ?

Question 2. Dans ce contexte, comment interprétez-vous l'arbre de droite ?

Votre encadrant vous confie six tubes de solution d'ADN génomique purifié provenant de personnes de la famille dont l'arbre est représenté en Figure 1, à droite, ainsi qu'un tube provenant d'une personne atteinte de DLFT/ALS sans lien de parenté et diagnostiquée par ailleurs.

Vous allez procéder à l'analyse de ces ADN par Southern blot. Avant de commencer toute analyse, vous devez utiliser une centrifugeuse afin de collecter les solutions au fond des tubes. Une centrifugeuse possède un rotor tournant à grande vitesse, tel celui présenté en Figure 2, muni de logements en périphérie recevant les tubes à centrifuger. Les tubes doivent être placés d'une manière équilibrée par rapport à l'axe central du rotor, sinon l'appareil risque d'être endommagé.



Figure 2 : Un rotor de centrifugeuse de laboratoire

Question 3. Comment allez-vous disposer les 7 tubes (qui sont tous identiques en masse) pour équilibrer le rotor ?

L'analyse de l'ADN par Southern blot se déroule globalement selon le protocole suivant :

1. L'ADN à analyser est digéré par des enzymes de restriction. Ces enzymes reconnaissent l'ADN au niveau de séquences spécifiques et coupent les deux brins.
2. Les fragments d'ADN obtenus sont déposés dans les puits à l'extrémité d'un gel d'électrophorèse qui est placé dans une cuve horizontale contenant une solution tampon et deux électrodes (+ et -) connectées à une source de courant.
3. Après migration, l'ADN est dénaturé, puis transféré sur une membrane placée sur le gel.
4. La membrane est ensuite mise en contact avec une sonde spécifique de la séquence d'ADN recherchée, ici le gène C9ORF72 (la sonde est elle-même un fragment d'ADN simple brin). La sonde a auparavant été marquée radioactivement ou par ajout de molécules fluorescentes.
5. Après lavage de l'excès de sonde, le gel est autoradiographié ou photographié.

Question 4. A quoi sert la digestion à l'étape 1 ?

Question 5. Comment doit-on brancher les pôles + et - du générateur ? Faites un schéma simple du montage et justifiez votre réponse.

Question 6. Sur quel phénomène repose l'utilisation de la sonde ?

La Figure 3 représente deux « Southern blots ». Le panneau A représente celui que vous avez obtenu à partir d'échantillons provenant directement des patients de l'arbre de la Figure 1, à droite (les numéros correspondent). Le panneau B représente le résultat d'un Southern blot conduit de la même façon, mais utilisant de l'ADN de cellules en culture. Ces cultures sont dites clonales. Cela signifie que, pour chaque échantillon, une cellule d'un patient a été isolée et cultivée : toutes les cellules d'une culture proviennent de la même cellule et forment donc un clone.

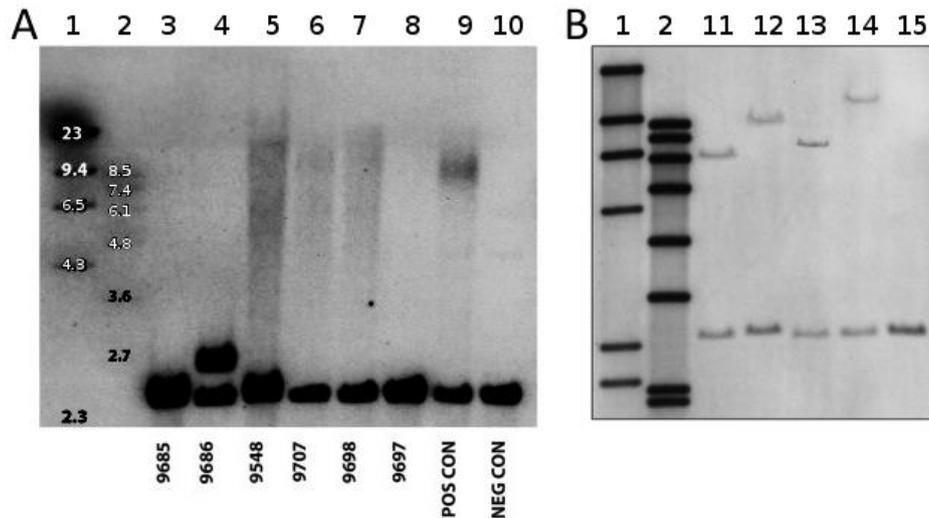


Figure 3 : Etude de l'ADN de patients et de sujets sains par Southern blot. A) Piste 1 marqueur de poids moléculaire avec des fragments de 2 ; 2,3 ; 4,3 ; 6,5 ; 9,4 et 23 kilopaires de bases (Kpdb). Piste 2 marqueur de poids moléculaire avec des fragments de 1,8 ; 1,9 ; 2,7 ; 3,6 ; 4,8 ; 6,1 ; 7,4 et 8,5 Kpdb. Pistes 3 à 8 : ADN des personnes de la famille étudiée, correspondent aux numéros dans la Figure 1, partie droite. Piste 9 (POS CON) : ADN de patient DLFT/ALS d'une autre famille et diagnostiqué par ailleurs. Piste 10 : (NEG CON) sujet non atteint. B) Piste 1 et 2 comme en A) ; Clones cellulaires créés à partir de cellules de personnes malades (piste 11 à 14) ou bien d'une personne saine (piste 15). Tirée de (3) et (2).

Question 7. Décrivez les résultats obtenus pour les ADN de patients.

Question 8. Quelles hypothèses pouvez-vous bâtir grâce à vos résultats ?

Voici la partie de la séquence du gène sauvage C9ORF72 (ici d'un sujet non atteint) dans laquelle surviennent les mutations:

```
5' CAGCCTGTAGCAAGCTCTGGAACCTCAGGAGTCGCGCCTAGGGGCCGGGGCCGGGGCCGGGGCGTGGTTCG
GGGCGGGCCCCGGGGCGGGCCCCGGGGCGGGGCTGCGGTTGCGGTGCCTGCGCCCCGGGCGGGCGGAGGCGC
AGGCGGTGGCGAGTGGGTGAGTGAGGAGGCGGCATCCTGGCGGGTGGCTGTTTGGGGTTCGGCTGCCGGG
AAGAGGCGCGGGTAGAAGCGGGGGCTCTCCTCAGAGCTCGACGATTTTTACTTTCCCTCTCATTTCTCT
GACCGAAGCTGGGTGTCGGGCTTTTCGCTCTAGCGACTGGTGAATTGCCTGCATCCGGGCCCGGGCTT
CCCCGGCGGGCGGGCGGGCGGGCGGCAGGGACAAGGGATGGGGATCTGGCCTCTTCCCTTGCCTTCCC 3'
```

Pour préciser la nature des anomalies détectées par Southern blot, une technique complémentaire appelée RP-PCR est souvent mise en œuvre dans ce contexte. Elle est basée sur la technique de PCR classique dont le principe est rappelé en annexe.

La RP-PCR (à ne pas confondre avec la RT-PCR) utilise 3 amorces appelées ici Forward 1 et 2 et Reverse. Les séquences sont données ci-dessous. Seules les amorces Forward 1 et Reverse peuvent trouver une cible dans l'ADN génomique. La réaction s'effectue avec une quantité d'amorce Forward 1 très limitée (par rapport aux deux autres amorces) et qui s'épuise rapidement dans les premiers cycles de la PCR.

```
Forward 1: 5' TACGCATCCCAGTTTGAGACGGGGGCCGGGGCCGGGGCCGGGG 3'
Forward 2: 5' TACGCATCCCAGTTTGAGACG 3'
Reverse : 5' AGTCGCTAGAGGCGAAAGC 3'
```

Question 9. Où l'amorce Reverse s'hybride-t-elle ?

Question 10. Que remarquez-vous à propos des deux amorces Forward ?

Question 11. Pouvez-vous schématiser le déroulement du premier cycle de la RP-PCR ?

Question 12. Que se passe-t-il dans les cycles suivants ?

La Figure 4 rapporte les résultats de RP-PCR chez un patient et un sujet sain. Le produit de la réaction a été analysé par électrophorèse capillaire capable de déterminer avec une résolution d'une paire de base la taille des fragments de PCR par ailleurs rendus fluorescents.

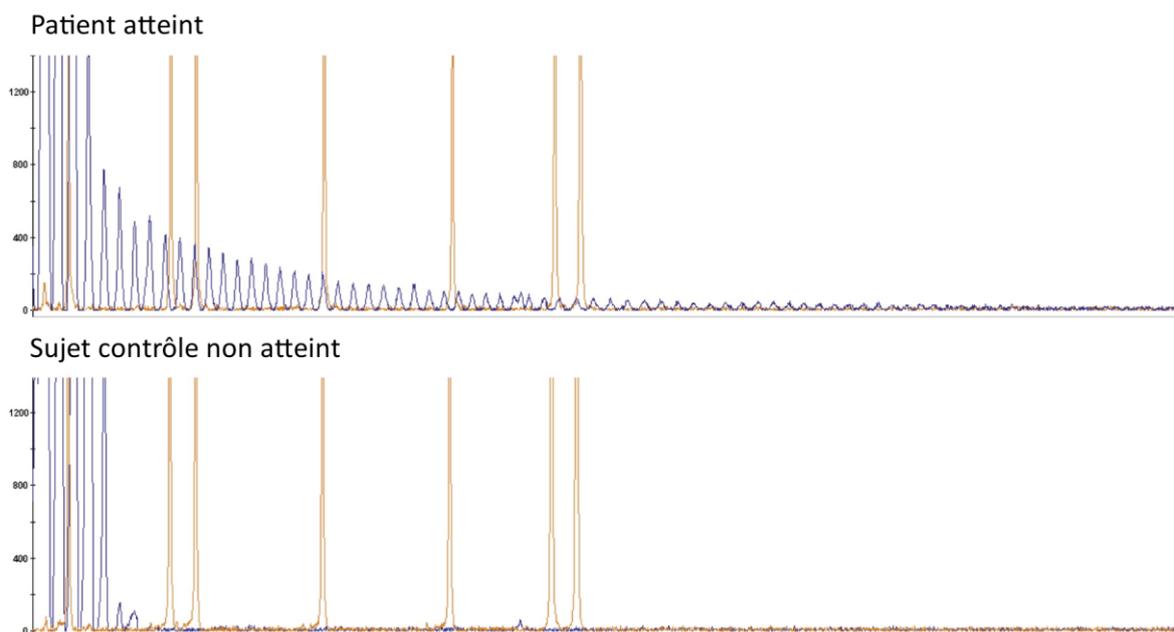


Figure 4. Résultats de RP-PCR chez un patient et un sujet contrôle. En ordonnées sont rapportées les mesures (en unités arbitraires de fluorescence) réalisées par l'appareil de chromatographie capillaire, et en abscisse, la longueur des fragments analysés. Les pics rouges sont des marqueurs de poids moléculaires. L'article dont est tirée la figure ne précise pas la taille de ces fragments. Les auteurs indiquent cependant que deux pics bleus consécutifs sont séparés de 6 paires de bases.

Question 13. Comment expliquer ces résultats ?

Un article récent rapporte les effets de l'inactivation du gène *C9orf72* chez la souris. La Figure 5 montre l'effet de l'inactivation sur le niveau d'expression du gène mesuré *via* l'ARN messager. La figure montre aussi un western blot. Pour cette technique, des extraits protéiques sont tout d'abord soumis à une électrophorèse en gel, puis les protéines sont transférées sur une membrane. Les protéines d'intérêt sont alors détectées sur la membrane grâce à des anticorps spécifiques.

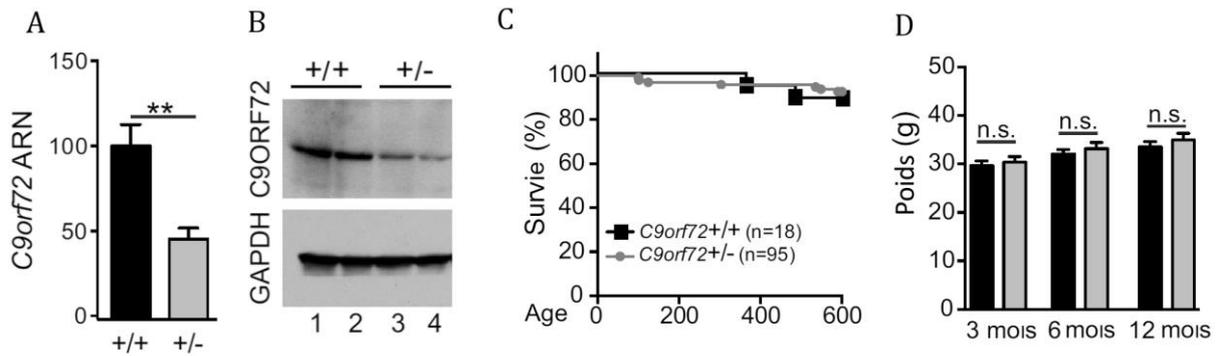


Figure 5. A. Expression du gène *C9orf72* chez la souris sauvage (+/+) ou hétérozygote (+/-) pour l'inactivation du gène. L'histogramme rapporte l'expression mesurée au niveau de l'ARN, et exprimée en pourcentage de l'expression chez la souris sauvage. B. Western blot dans lequel un anticorps dirigé contre la protéine C9ORF72 (panneau du haut) ou contre la GAPDH (Glyceraldéhyde-3-phosphate déshydrogénase, utilisée comme témoin) sont utilisés sur des protéines extraites de cerveaux de souris +/+ et +/- . C. Courbes de survie en fonction du temps des souris mutantes et sauvages. D. Evolution du poids des souris hétérozygotes et sauvages. ** : Différence statistiquement significative, n.s. : pas de différence.

Question 14. Que conclure de l'analyse de cette figure ?

Le labyrinthe de Barnes est un outil de mesure de la mémoire de référence spatiale chez les souris et les rats. Il permet d'estimer la capacité de l'animal à se repérer grâce à des indices spatiaux, situés dans la pièce, pour trouver la bonne sortie d'un labyrinthe. Le test est répété chaque jour pendant 9 jours, et les résultats sont exprimés en nombre moyen d'erreurs sur 3 jours. La Figure 6 rapporte les résultats.

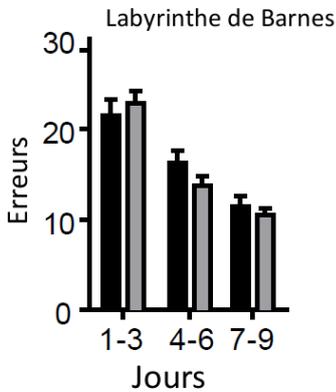


Figure 6. Performance des souris sauvages (barres noires) ou mutantes hétérozygotes pour *C9orf72* (barres grises) dans le test du labyrinthe de Barnes.

Question 15. Que conclure de l'analyse de cette figure ?

Les auteurs ont répété ce type de test avec des souris de deux lignées transgéniques différentes qui cette fois portent une copie d'un allèle mutant de *C9ORF72* humain isolé de patients ; chaque lignée contient un allèle différent.

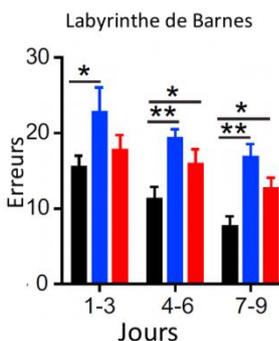


Figure 7. Performance des souris sauvages (barres noires) ou portant un allèle mutant humain de *C9ORF72* (barre bleues et rouges) dans le test du labyrinthe de Barnes. * et ** indiquent des différences significatives. Adapté de (4).

Question 16. Quelles conclusions tirer de l'analyse de cette figure ?

Question 17. Proposer une hypothèse pour expliquer la différence de résultats entre les deux lignées.

Afin de mettre en évidence les conséquences des mutations de C9ORF72, plusieurs laboratoires ont développé des anticorps pour détecter les protéines suivantes : poly (GA), poly (GP) et poly (GR). Poly(GA) par exemple désigne une protéine contenant la séquence Glycine-Alanine-Glycine-Alanine-etc... répétée un certain nombre de fois. La figure 8 représente le résultat d'une détection par immunohistochimie avec l'anticorps anti poly (GA) sur des neurones de cortex frontal de patients DLFT/ALS. La présence de la protéine recherchée est signalée par l'apparition d'une couleur brune. Tous les noyaux cellulaires sont colorés en bleu. Les 3 anticorps sont capables de détecter les protéines chez les patients, mais pas chez les sujets en bonne santé.

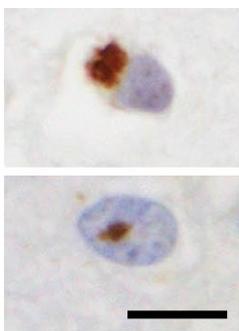


Figure 8. Immunodetection de protéine poly (GA) sur des neurones après autopsie. Barre d'échelle 20 μ m. Adaptée de (5).

Question 18. Où situez-vous le marquage obtenu ? Quelles sont les fonctions de ces structures ?

Question 19. Pourquoi les laboratoires ont-ils développé ces anticorps en particulier, plutôt que n'importe quel autre ?

Question 20. En vous rappelant que poly (GP) et poly (GR) sont aussi détectées, que vous suggère l'observation de la Figure 8 ?

Question 21. Il est également possible de détecter, à un niveau plus faible, l'expression des protéines poly (PR) et poly (AP). Qu'implique cette observation ?

Un laboratoire a utilisé la séquence $(GGCCCC)_4$, c'est-à-dire un oligonucléotide de 24 bases, rendu fluorescent, pour réaliser une expérience dite d'hybridation *in situ* sur des noyaux de neurones du cortex ou de la moëlle épinière (spinal cord) de sujets malades ou de témoin (contrôle). Ils ont aussi utilisé une sonde témoins $(CAGG)_4$.

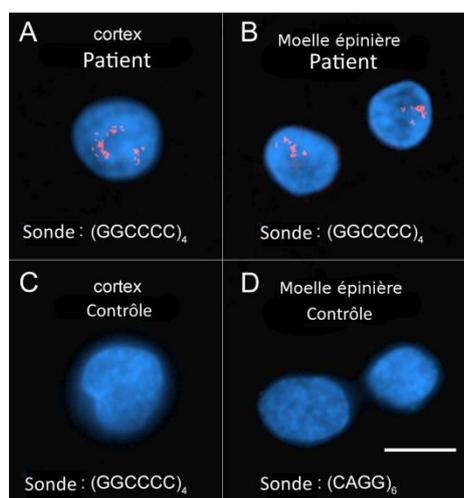


Figure 9. Hybridation *in situ* sur des noyaux de neurones de sujets malades DLFT/ALS ou de sujets sains. Les sondes apparaissent en rouge et l'ADN, coloré par le DAPI, apparaît bleu. Barre d'échelle : 20 μ m. Adaptée de (3).

Question 22. Que les sondes détectent elles ?

Question 23. Que donneraient des sondes de séquence GGGGCC ?

Question 24. Quels impacts sur le fonctionnement cellulaire peut avoir la présence dans le noyau des molécules détectées ?

Question 25. Pouvez-vous synthétiser dans un schéma l'ensemble des résultats de cette première partie ?

Partie II 1h.

Fonction de H2A.Z dans la mémoire

Le fonctionnement de la mémoire repose sur des mécanismes très divers. Une étude récente (6) s'est intéressée aux rôles joués par la protéine histone H2A.Z.

Les histones sont des protéines chargées positivement à pH physiologique. Elles s'assemblent en octamères qui, en s'associant avec l'ADN, forme un complexe appelé nucléosome. Chaque nucléosome contient environ 150 paires de bases d'ADN, et constitue ainsi le niveau le plus bas de l'organisation de l'ADN sous forme de chromatine dans le noyau. Chaque octamère est constitué de deux exemplaires de chacune des 4 histones dites canoniques (appelées H2A, H2B, H3 et H4). Certains nucléosomes voient une ou plusieurs histones canoniques être remplacées par une histone dite « variante ». H2A.Z est l'une de ces variantes.

Question 26. Que signifie exactement « protéines chargées positivement à pH physiologique » ? Comment évolue leur charge en fonction du pH ?

Les auteurs ont mis en œuvre le protocole de conditionnement à la peur suivant. Des souris ont été placées dans une chambre d'entraînement pendant 2 min pour qu'elles l'explorent puis ont reçu 3 chocs électriques aux pattes ; une minute leur est donnée ensuite pour explorer la chambre puis elles sont retirées. Des souris naïves (N) non entraînées ont été utilisées comme témoins. Les animaux ont été mis à mort 30 min ou 2 h après ce conditionnement et les cerveaux ont été prélevés.

Les auteurs se sont intéressés aux nucléosomes placés immédiatement en amont (appelés -1 nucléosome) et en aval (+1 nucléosome) du point de départ de la transcription (appelé TSS, transcription start site) de certains gènes exprimés dans la région CA1 de l'hippocampe du cerveau, zone importante pour la mémoire. Ils ont aussi mesuré l'expression de ces gènes (mesure de la quantité d'ARNm, notée mRNA). La technique de CHIP (immunoprécipitation de la chromatine) permet de mesurer la présence de H2A.Z et de l'exprimer en pourcentage par rapport au groupe témoin (Souris naïves n'ayant subi aucun conditionnement). Les gènes choisis sont connus pour être soit des activateurs (*Npas4*, *Egr2* et *Arc*) soit un répresseur de la mémoire (*Ppp3ca*). Les résultats sont synthétisés dans la Figure 10.

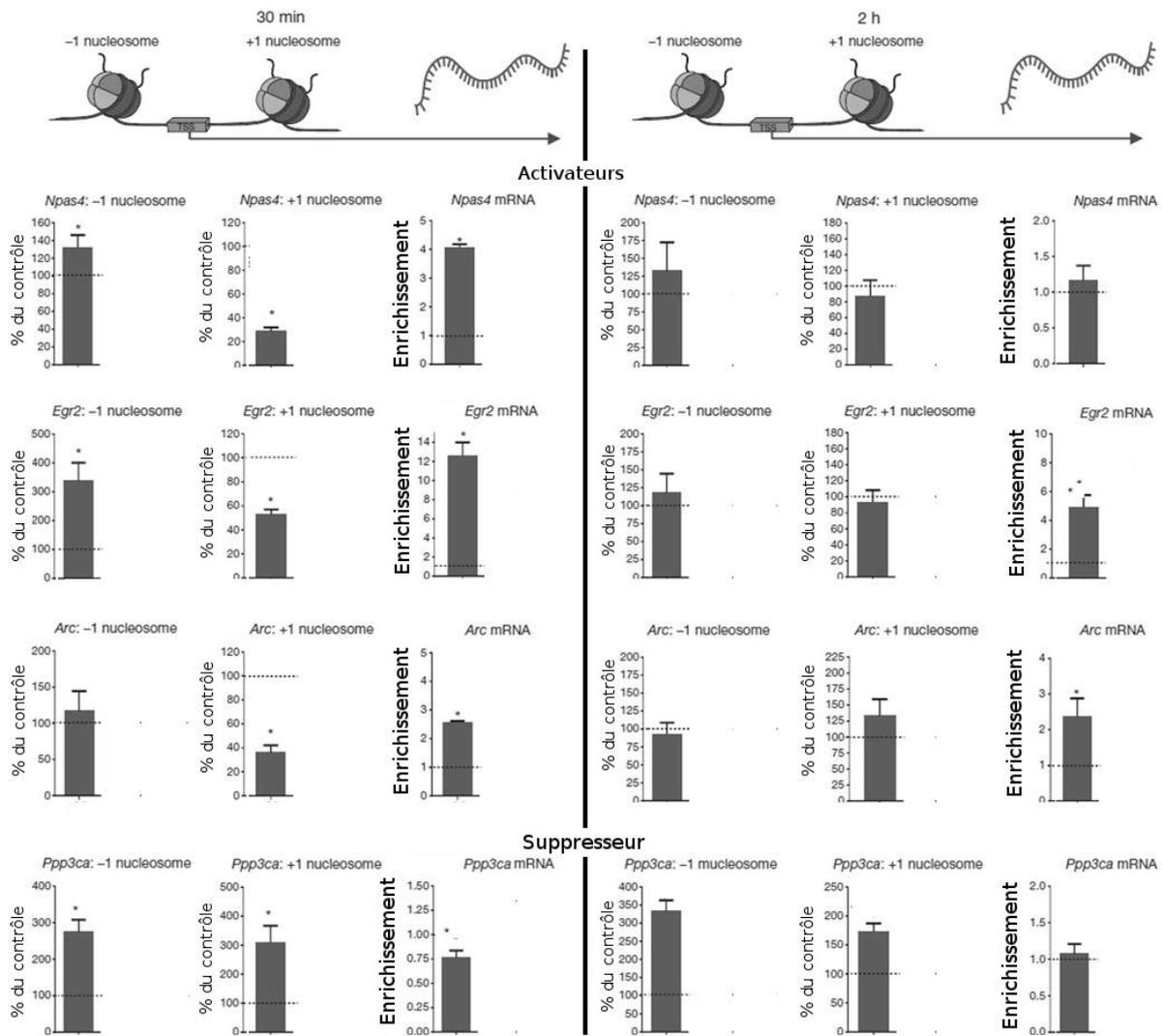


Figure 10. Présence de l'histone H2A.Z autour du point de départ de la transcription de gènes activateurs ou supresseur de la mémoire dans l'hippocampe, et transcription de ces gènes 30 minutes (histogrammes les plus à gauches) ou 2 heures (les plus à droites) après conditionnement. * : différence significative.

Question 27. Commenter cette figure.

Question 28. A priori, quelle est l'influence globale de l'histone H2A.Z sur la mémoire ?

Il est possible d'évaluer le souvenir que les souris gardent de leur conditionnement. Pour cela, elles sont replacées dans la chambre de conditionnement et observées. Les souris qui se souviennent que la chambre est associée aux chocs électriques s'immobilisent comme figées pendant un temps plus ou moins long (comportement appelé « Freezing »). Les auteurs ont évalué ainsi la mémoire de souris qui ont de plus été traitées par injection dans l'hippocampe soit d'un virus (dérivé du « Adeno-associated Virus », AAV) capable de diminuer l'expression de H2A.Z (H2A.Z AAV) soit d'un virus contrôle (Scr AAV).

Question 29. Avez-vous une idée de la façon dont on peut ainsi interférer avec l'expression d'un gène particulier ?

Les résultats de l'expérience avec les virus sont rapportés en Figure 11. Après injection des AAV et deux semaines d'incubation (soit J14), les animaux sont conditionnés comme décrit plus haut, puis testés le lendemain ou deux semaines plus tard. Le niveau d'expression de H2A.Z est aussi mesuré à J14.

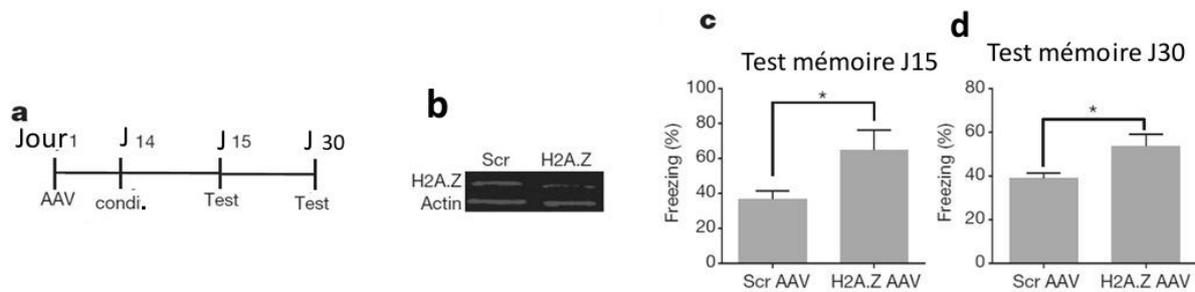


Figure 11. Test de la mémoire conditionnée à la peur. a) Protocole utilisé. b, western blot avec des anticorps anti H2A.Z et anti actine (utilisée comme contrôle) à J14. c et d) Effet du traitement par Scr AAV ou par H2A.Z AAV. Pour une description de la technique de western blot, voir en fin de page 5.

Question 30. Ces nouvelles expériences confirment-elles les conclusions précédentes ?

Question 31. Pouvez-vous synthétiser dans un schéma l'ensemble des résultats de cette partie ?

Question 32. Pourriez-vous proposer une question supplémentaire que les auteurs auraient pu poser?

Question 33. Pouvez-vous imaginer des mécanismes expliquant les propriétés observées de H2A.Z ?

Références utilisées:

1. A. E. Renton *et al.*, *Neuron*. **72**, 257–268 (2011).
2. Z. Xi *et al.*, *Am. J. Hum. Genet.* **96**, 962–970 (2015).
3. M. DeJesus-Hernandez *et al.*, *Neuron*. **72**, 245–256 (2011).
4. J. Jiang *et al.*, *Neuron*. **90**, 535–550 (2016).
5. I. R. A. Mackenzie *et al.*, *Acta Neuropathol. (Berl.)*. **130**, 845–861 (2015).
6. I. B. Zovkic, B. S. Paulukaitis, J. J. Day, D. M. Etikala, J. D. Sweatt, *Nature*. **515**, 582–586 (2014).

Annexes :

Le code génétique :

		Deuxième lettre				
		U	C	A	G	
Première lettre	U	UUU } Phe UUC } UUA } Leu UUG }	UCU } UCC } Ser UCA } UCG }	UAU } Tyr UAC } UAA Stop UAG Stop	UGU } Cys UGC } UGA Stop UGG Trp	U C A G
	C	CUU } CUC } Leu CUA } CUG }	CCU } CCC } Pro CCA } CCG }	CAU } His CAC } CAA } Gln CAG }	CGU } CGC } Arg CGA } CGG }	U C A G
	A	AUU } AUC } Ile AUA } AUG Met	ACU } ACC } Thr ACA } ACG }	AAU } Asn AAC } AAA } Lys AAG }	AGU } Ser AGC } AGA } Arg AGG }	U C A G
	G	GUU } GUC } Val GUA } GUG }	GCU } GCC } Ala GCA } GCG }	GAU } Asp GAC } GAA } Glu GAG }	GGU } GGC } Gly GGA } GGG }	U C A G
		Troisième lettre				

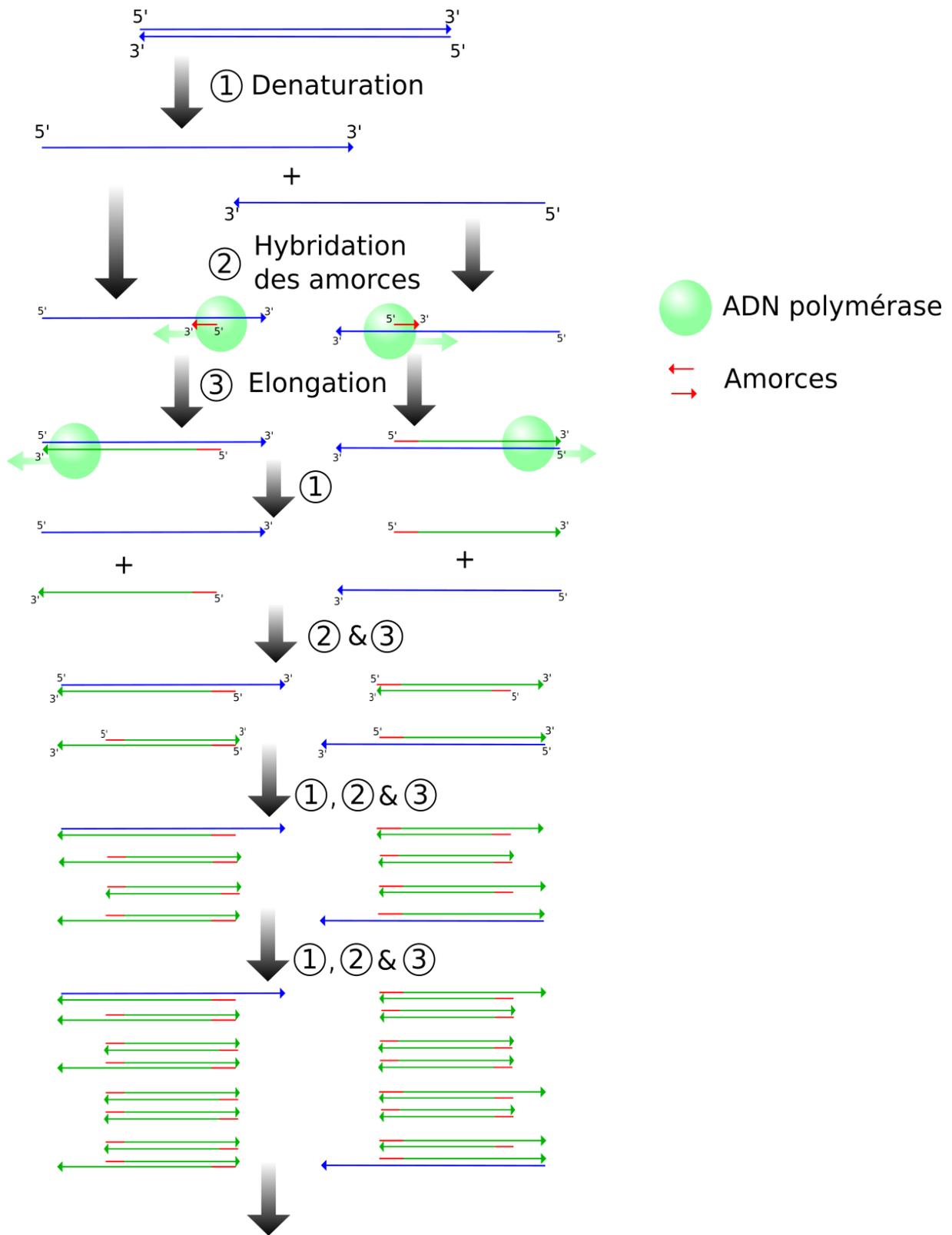
Les abréviations utilisées pour les acides aminés :

Acide aminé	Code à 3 lettres	Code à 1 lettre
alanine	ala	A
arginine	arg	R
asparagine	asn	N
acide aspartique	asp	D
asparagine ou acide aspartique	asx	B
cystéine	cys	C
acide glutamique	glu	E
glutamine	gln	Q
glutamine or acide glutamique	glx	Z
glycine	gly	G
histidine	his	H
isoleucine	ile	I
leucine	leu	L
lysine	lys	K
méthionine	met	M
phénylalanine	phe	F
proline	pro	P
serine	ser	S
thréonine	thr	T
tryptophane	trp	W
tyrosine	tyr	Y
valine	val	V

Le principe de la PCR (voir aussi schéma page suivante)

La technique repose sur l'extension d'un couple d'amorces (typiquement des oligodésoxynucléotides longs d'une vingtaine de bases) dont la séquence est choisie pour pouvoir s'hybrider par complémentarité des bases de part et d'autre de la région que l'on souhaite copier (la matrice). La réaction est catalysée par une ADN polymérase thermostable en présence de désoxynucléotides triphosphates.

En pratique, la réaction se déroule dans un thermocycleur, appareil pouvant faire évoluer la température du milieu réactionnel rapidement et précisément. L'ADN à copier (par exemple, de l'ADN génomique) est dénaturé par augmentation de la température. Les amorces s'hybrident sur la matrice, en flanquant la région à copier, à la faveur d'une baisse de température et la polymérase synthétise les brins complémentaires en allongeant les amorces. On fait succéder à ce premier cycle dénaturation-hybridation-élongation autant de cycles que souhaités. Les portions nouvellement synthétisées servent ensuite de matrice, ce qui confère un caractère exponentiel à la réaction.



Augmentation exponentielle du nombre de copies de la séquence.