

Tous les énoncés des questions suivantes peuvent servir à répondre aux questions 15 à 42.

A la fin du fascicule, les trois séquences correspondent soit à des séquences d'ADNc complètes soit à des séquences d'ADN génomique. L'une concerne la protéine SRY, les deux autres la protéine SOX9. Tous les exons contiennent une séquence codante. Ces protéines ne sont pas sécrétées, mais sont nucléaires. SRY permet la différenciation de la gonade bipotentielle en testicule. SOX9 participe à la même action, mais a un rôle primordial sur la biosynthèse des collagènes.

QUESTION N° 15 (2 points)

En sachant que la partie 3' non codante de l'ADNc (queue polyA exclue) de SRY est de 134 pb (à 3 pb près), on peut dire que

- A la séquence 2 est un ADNc
- B les séquences 2 et 3 sont des séquences génomiques
- C la séquence 1 est l'ADN génomique de SOX9
- D la séquence 2 est l'ADNc de SRY
- E la séquence 3 est l'ADNc de SOX9

QUESTION N° 16 (2 points)

Le gène *SRY* « Sex-determining Region of the chromosome Y » a été découvert en 1990 après des années de recherche. La protéine SRY contient un domaine HMG « High Mobility Group » constitué de 79 acides aminés (le premier correspond à Asp58). Toutes les protéines possédant un domaine HMG, dont l'homologie de séquence est égale ou supérieure à 60% avec celle de SRY, appartiennent à la famille des protéines SOX (SRY-related high mobility group bOX). La séquence de la protéine humaine SOX9 a été découverte en 1993 (année où la séquence du génome humain n'était pas connue). D'après vous, quelles sont les étapes suivantes les plus logiques et les plus efficaces qui ont dû être utilisées pour déterminer la séquence de la protéine SOX9?

- A Purification de la protéine et séquençage protéique
- B Méthodes basées sur l'homologie de séquences avec le domaine HMG de SRY
- C Criblage d'une banque complémentaire avec un fragment d'ADNc de 230 pb commençant au nucléotide 520 de la séquence 2
- D Utilisation de colonne d'affinité avec un anticorps fait à partir d'un peptide constitué des acides aminés 160 à 213 de la protéine SRY
- E Criblage d'une banque complémentaire avec le fragment d'ADNc qui code le domaine HMG de SRY

QUESTION N° 17 (3 points)

En examinant les séquences, on peut dire que

- A le domaine HMG de la protéine SOX9 est codé par un exon
- B la protéine SOX9 a 510 acides aminés
- C le gène *SOX9* contient 2 introns
- D le transcrit primaire du gène *SOX9* est long de 5398 nucléotides (à 3 nucléotides près)
- E la protéine SOX9 contrairement à SRY contient un domaine codé par l'exon 3 constitué uniquement de trois acides aminés : proline, alanine et glutamine

QUESTION N° 18 (1 point)

En examinant les séquences, on peut dire que

- A la transcription du gène *SRY* commence au nucléotide 352 (à 3 nucléotides près)
- B la protéine SRY contient 204 acides aminés
- C le gène *SRY* ne contient pas d'intron
- D le domaine HMG de la protéine SRY est codé par un exon
- E le gène *SRY9* a une séquence de signal de polyadénylation : AATAAA

QUESTION N° 19 (2 points)

En sachant qu'un seul des changements du codon 169 du gène *SOX9* n'est pas pathogène, on peut dire que

- A la mutation c.507C>A est pathogène et s'écrit aussi p.His169His
- B le changement du nucléotide 506 peut donner une mutation silencieuse
- C le changement du nucléotide 505 C en T n'est pas pathogène
- D le changement du nucléotide 505 C en T peut donner un SNP
- E la mutation c.507C>T peut s'écrire p.His169Gln

QUESTION N° 20 (1 point)

Une pathologie liée à SOX9 se transmet selon un mode autosomique dominant. Généralement à propos de ce mode de transmission, on peut dire que

- A la distribution des cas est verticale sur un arbre généalogique
- B la maladie touche les deux sexes
- C un individu malade est hétérozygote ou homozygote pour l'allèle muté
- D un individu sain aura ses deux parents sains, si la pénétrance est complète
- E le cas d'un individu atteint avec des parents sains peut s'expliquer par la survenue d'une mutation de novo (néo-mutation)

QUESTION 21 (1 point)

Un fœtus 46,XY atteint d'un syndrome campomélique (anomalies du squelette importantes avec agénésie des omoplates, os longs courts, fémurs incurvés,...) et ayant un phénotype féminin est suspecté d'avoir une lésion du gène *SOX9*. Il s'agit d'une maladie génétique très grave entraînant souvent la mort soit in utero soit à la naissance par insuffisance respiratoire (cage thoracique étroite et déformée). Le plus souvent les parents ne sont pas atteints. Pour rechercher une mutation du gène *SOX9*,

- A une ponction du liquide amniotique a été effectuée.
- B une prise de sang des deux parents a été faite pour savoir si l'un des parents a transmis la mutation
- C l'extraction de l'ADN a été faite sur le surnageant du liquide amniotique après l'avoir centrifugé
- D une culture des amniocytes a été faite pour extraire l'ADN
- E le dosage de la protéine *SOX9* a été réalisé sur le surnageant du liquide amniotique après l'avoir centrifugé

QUESTION 22 (1 point)

Pour séquencer le gène *SOX9* en entier, plusieurs fragments d'ADN ont été amplifiés. L'un d'entre eux est long de 640 pb et permet d'amplifier un exon. L'amorce 1 commence par le nucléotide 2001.

Quelles sont les amorces 1 et 2 qui ont été utilisées ?

- A 5' TTCTTGCAGCTTAAGTAAT 3'
- B 5' GGGTGCTTCTTAGAGGGTCC 3'
- C 5' TTAGGGGGCAGCACCGGGAG 3'
- D 5' CCCACGAAGAATCTCCCAGG 3'
- E 5' GGACCCTCTAAGAAGCACCC 3'

QUESTION 23 (3 points)

Le fragment amplifié a été séquencé en utilisant les mêmes amorces qui ont servi à l'amplifier (amorces 1 et 2). Une mutation a été identifiée dans ce fragment d'ADN génomique en lisant les séquences obtenues. Ci-dessous une partie de la séquence obtenue englobant la fin de l'exon de l'ADN du fœtus. La lettre N signe l'impossibilité du logiciel du séquenceur de déterminer le nucléotide. Le plus souvent, c'est le signe d'une double séquence. Quelle que soit l'amorce utilisée, le logiciel donne toujours les séquences 5'-3' que vous avez à la fin du fascicule.

Séquence A

5'...NCCCCNGNNNNNNNNNGGNGAGTCGCCCCTCGACCCCA.....3'

Séquence B

5'...CTCCCCGGCGAGCACTCGGGNNNNNNNCCCNNNNNCCC.....3'

Après cette expérience et la lecture des séquences, on peut dire que

- A le fœtus est homozygote pour cette mutation
- B la séquence A est obtenue en utilisant l'amorce 2
- C la mutation est une mutation ponctuelle changeant un acide aminé
- D la mutation est une délétion d'un nucléotide
- E la mutation est située dans l'exon

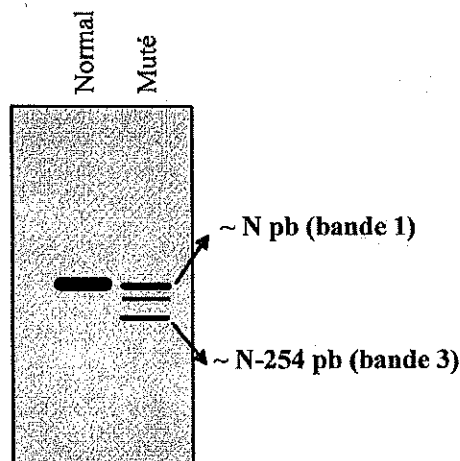
QUESTION 24 (1 point, 1 seule réponse juste)

Le nom de la mutation

- A c.685delG
- B c.687delG
- C p.Gln230Asnfs23
- D IVS2+1G>T
- E IVS2+2delT

QUESTION 25 (1 point)

Comme cette maladie entraîne la mort soit in utero soit rapidement à la naissance, une interruption de grossesse est faite. Il a été possible pour déterminer le retentissement de cette mutation d'utiliser le testicule pour réaliser une RT-PCR avec pour l'étape PCR l'utilisation de deux amorces, l'une située dans l'exon 1 et l'autre dans l'exon 3. Trois fragments sont obtenus après l'électrophorèse des produits de la RT-PCR (à 10 pb près) (ci-dessous la photo de l'électrophorèse après révélation).



Électrophorèse de produits de RT-PCR

- A la RT-PCR a été réalisé à partir de l'ADN extrait du tissu testiculaire
- B la photo correspond à celle d'une autoradiographie
- C les bandes ont été détectées aux UV après coloration avec le bleu de Coomassie
- D le fragment N-254 pb ne contient pas l'exon 2
- E le fragment de N pb correspond à l'expression de l'allèle normal

QUESTION 26 (2 points)

La lecture du séquençage du fragment de RT-PCR, correspondant à la bande intermédiaire située entre la bande 1 et la bande 3, permet de définir l'une des conséquences de cette mutation sur l'ARNm.

Ci-dessous une partie de la séquence d'intérêt (début vers le nucléotide 2500) :

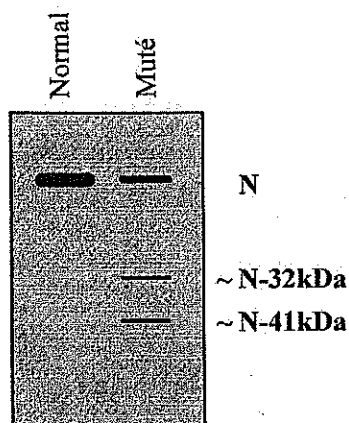
5'...ATGAGCGAGGGCAATCCCAGGGCCCA...3'

On peut conclure d'après les données ci-dessus et celles de la question précédente que cette mutation

- A crée un nouveau site accepteur
- B ne modifie pas l'épissage
- C donne une protéine SOX9 plus longue
- D donne deux protéines tronquées dont les 229 premiers acides aminés sont identiques à la protéine SOX9
- E entraîne un exon skipping de tout l'exon 2

Enoncé pour les deux questions suivantes

Pour compléter le retentissement de cette mutation sur la protéine SOX9, un Western blot a été réalisé à partir des extraits testiculaires. Le résultat est obtenu grâce à un anticorps anti-immunoglobuline de souris marqué à l'iode 125. Les valeurs en kDa sont données à 2kDa près.



Western blot des extraits testiculaires

QUESTION N° 27 (1 point)

À propos de la réalisation de ce western blot,

- A Les bandes vues sur ce western blot sont colorées par le bleu de Coomassie
- B L'anticorps anti-SOX9 est un anticorps monoclonal reconnaissant l'extrémité N-terminale de la protéine
- C L'anticorps anti-SOX9 a été incubé après la séparation des extraits testiculaires sur gel SDS-acrylamide et avant le transfert sur nylon
- D L'anticorps anti-SOX9 est un anticorps polyclonal reconnaissant l'extrémité C-terminale de la protéine
- E l'image ci-dessus est celle d'une autoradiographie.

QUESTION N° 28 (2 points)

A propos des résultats précédents et de ce western blot, on peut dire que

- A la valeur de N est d'environ 60000 daltons
- B une seule des deux protéines tronquées contient la boîte HMG
- C la protéine N-41 kDA est produite à la suite de la création d'un site donneur dans l'exon 2
- D la protéine N-32 kDA contient la boîte HMG
- E l'atteinte du fœtus est due à une haploinsuffisance c'est-à-dire qu'il n'y a pas assez de protéines SOX9 pour stimuler la biosynthèse des différentes chaînes de collagène et pour former un testicule fonctionnel.

QUESTION N° 29 (2 points)

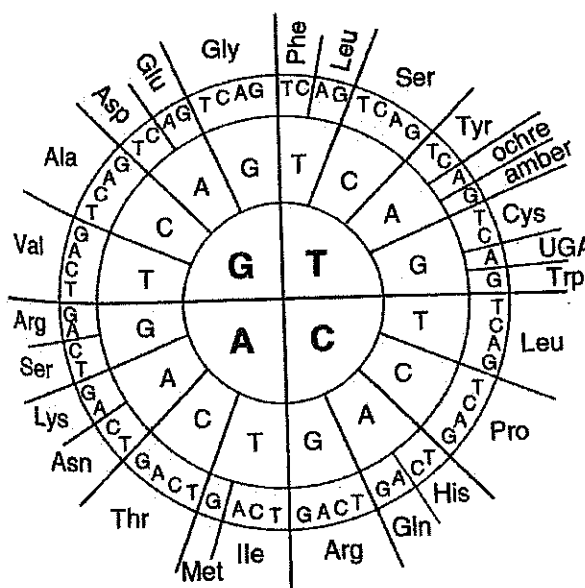
La séquence 5' UCACAUAAACACAAUUCUGUU 3' qui est complémentaire au fragment contenant le nucléotide 4630 pourrait être

- A une amorce de PCR
- B un micro ARN
- C un fragment d'ARNm
- D un ADN interférent
- E une amorce de RT-PCR

Séquence 1

GAGCATGTTAATCTATTTATATGGATTATTACGGAGGAACAGCGGGCGTTGAGTCACCAAACATTTGCTTCAAAGACT 80
 ATTTCTAAGCACTTTTGCAGGCAGGCAGGCTCGCTCCAGGCGCGTAAACTCGGCTACGCATTAAGAAGCGGCTGCTTTTC 160
 GAATACTGCAAACCTCCAGCTAAGTCCCCGGTCCCGCGGAGAGAGCAGTGAAAAGAAATGTCGGAGGTGGGGTAGATCCCT 240
 AGTCTAGACACACACTTGCAGCACACACACACACACACACACAAAGATTCGCGCGGAGAAGGCACATAAAATCTGGCA 320
 TTCCGAGAGTACGACAACTTACACACTTGGAAAGTCCCGGGTCCCCCGCTTCCCCCGCAGCACCCCCCGCCCCCACC 400
 TACCGTCCGCCCTTTGGCTGCGATCCCCTCCCCTCTCCTCCCCCTCCCGCTCGTCACCCAGCCAGTGCCACAATCCTCC 480
 TCCCTCCCCAAAATCGGGTCCAATCAGCTGCCTGCCAACCCCTGGGACTGCTGTGCTGTGATTGGCGGGTGGCTCTAAGGT 560
 GAGGCGGAGTATTTATTAAGAGACCCTGGGCTGGGAGTTGGAGAGCCGAAAGCGGAGCTCGAAACTGACTGGAACTTC 640
 AGTGGCGCGGAGACTCGCCAGTTTCAACCCCGGAACTTTTCTTTGCAGGAGAGAAGAGAAGGGGTGCAAGCGCCCCCA 720
 CTTTTGCTCTTTTTCTCCCCTCCTCCTCCTCCTCCAATTCGCCTCCCCCACTTGGAGCGGGCAGCTGTGAACCTGGCCAC 800
 CCGCGCCTTCCTAAGTGTGCTCGCGCGGTAGCCGGCCGACGCGCCAGCTTCCCGGGAGCCGCTTGTCTCCGCATCCGGGC 880
 AGCCGAGGGGAGAGGAGCCCGCCTCGAGTCCCCGAGCCGCGCGGCTTCTCGCCTTTCCCGGCCACCAGCCCCCTGCC 960
 CCGGGCCCGGTATGAATCTCCTGGACCCCTTCATGAAGATGACCGACGAGCAGGAGAAGGGCCTGTCCGGCGCCCCAG 1040
 M N L L D P F M K M T D E Q E K G L S G A P S 23
 CCCCACCATGTCCGAGGACTCCGCGGGCTCGCCCTGCCCTCGGGCTCCGGTCCGACACCGAGAACACGCGGCCCCAGG 1120
 P T M S E D S A G S P C P S G S G S D T E N T R P Q 50
 AGAACAGTTCCCCAAGGGCGAGCCCGATCTGAAGAAGGAGAGCGAGGAGGACAAGTTCCCCGTGTGCATCCGCGAGGCG 1200
 E N T F P K G E P D L K K E S E E D K F P V C I R E A 76
 GTCAGCCAGGTGCTCAAAGGCTACGACTGGACGCTGGTGCCATGCCGGTGCAGGTCACCGCTCCAGCAAGAACAAGCC 1280
 V S Q V L K G Y D W T L V P M P V R V N G S S K N K P 103
 GCACGTCAAGCGGCCATGAACGCCTTCATGGTGTGGGCGCAGGCGGCGCAGGAAGTTCGCGGACCAGTACCCGCACT 1360
 H V K R P M N A F M V W A Q A A R R K L A D Q Y P H 130
 TGCACAACGCCGAGCTCAGCAAGACGCTGGGCAAGCTCTGGAGGTAGGACCGGGCGGGGGCGGGCAGGGTGGGCAT 1440
 L H N A E L S K T L G K L W R 144
 CGCGCGGCTGAGGGCTGATTTGCCCGCCCGCCCTCCCATCGCCCGGGAGTTGCCGTTCCGGGAGCC 1520
 GGCGGGATGGGGTGGGAGTGGGAATGGGGTGTAACTGTGGCTCAGAGTTTGACAAAGTCTTGGGCTGCTCGCGGGGAC 1600
 GCGGAGGAGGGGGTGGTAAGTGGAAAGAGGTGAGGGAGGTAGCTGGAGGATGGACGAAGACTGGTGGGAGACGGAAGGAG 1680
 GGGGCTGCCAGCCTGCTCTCCAGTCGCCTGGAAGCTCAATCGGGGCGGGGAAGTGAACCTGCCTCCCTCCTACCCGGCC 1760
 TCTTAAAACCTGCACTCTCTCGTGCAGCCCCACTGTCCACGGAGATGGGGCAAGGGAGAAACCGAGGTGGAGGAGACCT 1840
 TGGCAGGAACTGGGAGGCGGGAGGAGGGAGGCTACTGGAATAGGTGGGAGTGTATGGTGGGGGTGAGAATTGGGGACC 1920
 TTCTTGCAGCTTAAGTAATTTGGGGAAAGTTTCAAAGGGGGTGGGGTGGGGCGGTAAGTCGAGCAGCAAAGGCGT 2000
 TTAGGGGGCAGCACCGGGAGTCGTTTTTCATCTCCAGCGTTTCAAATAGAAATAGAAGGGGAGGGGAGGGGGGGCGG 2080
 GGAGTGACCGCTCAGGTCAGACTGCAATAACTTATTTATTTATTTATTTTAAAGAAAAGTTATGAGCTGTGGTTGCAGGC 2160
 AGGAGGGAAGATGGAGTTGTGTGCAGAGGAAGCCGAGTGGTCTGGGTGCGCCCTCCTCCCCCGCCAGCTGACAGTTTGG 2240
 CGGATTTCACTGACCCCTCTCCCTCTTTTCTCTGTGCCCCCGCCCCGCCCCGAGCAGACTTCTGAACGAGAGCGAGAA 2320
 L L N E S E K 151
 GCGGCCCTTCGTGGAGGAGGCGGAGCGGCTGCGCGTGCAGCACAAGAAGGACCACCCGGATTACAAGTACCAGCCGCGG 2400
 R P F V E E A E R L R V Q H K K D H P D Y K Y Q P R 177
 GGAGGAAGTCCGGTGAAGAACGGGCAGGCGGAGGCAGAGGAGGCCACGGAGCAGACGCACATCTCCCCAACGCCATCTTC 2480
 R R K S V K N G Q A E A E E A T E Q T H I S P N A I F 204
 AAGGCGCTGCAGGCCGACTCGCCACACTCCTCCTCCGGCATGAGCGAGGTGCACTCCCCCGGCGAGCACTCGGGTGAAGTC 2560
 K A L Q A D S P H S S S G M S E V H S P G E H S 228
 GCCCTCGACCCACCGGACAAGCTATCTCGTCCCGCTGGCACACCCCTGCCCTCCGCTGGGAGATPCTTCGTGGG 2640
 GACTTTATGCTTCCCGGGAGGGACACACTGCCCTTTGCGCCCGTCCCGCTCCCTCCTTACCCAGAGCCTAAGAGGCATC 2720
 CAAACAACACACACAAACACACACACCCCAACTCAATCCCAGCATCCGAAGAGATTAACTTTTTTATTTGGGAGGTAAA 2800
 ATGCCCTTAACAGCCTTACAAGACCTCTCCCTTCTCTGCTCCCCACCCCAAAGCACACACAGGGCTCTTACACAA 2880
 GTAGCAATTAGGCTTCCGGACCTCCGGGCCAGACCCCTCCCTGATAAAAGGGGCTGTCCAGTGTGTACCGGCGGG 2960
 TTAATCATTGGGCGACTTATCTCCGGTGCAGCGCCCTCTTTCGCGGGTGCGGCCCTTATACACTTTAGCAGCGAGGG 3040
 AGGGTCCCCGGAGGGTGCCTAAGACTAGGGGCTCTGCACAGCCCTTGTGATTTTTCTCGTGTGTCTTTTATTGTCCA 3120
 CAGGGCAATCCCAGGGCCCACCGACCCACCCACCCCAAAAACCGACGTGCAGCCGGGCAAGGCTGACCTGAAGCGA 3200
 G Q S Q G P P T P P T T P K T D V Q P G K A D L K R 254
 GAGGGGCGCCCTTGCCAGAGGGGGCAGACAGCCCTATCGACTTCCGCGACGTGGACATCGGCGAGCTGAGCAGCGA 3280
 E G R P L P E G G R Q P P I D F R D V D I G E L S S D 281
 CGTCATCTCAAACATCGAGACCTTCGATGTCAACGAGTTTGACCAGTACCTGCCGCCAACGGCCACCCGGGGTGGCCG 3360
 V I S N I E T F D V N E F D Q Y L P P N G H P G V P 307
 CCAAGCAGGGCAGGTACCTACAGGGCAGCTACGGCATCAGCAGCACCGCGGCCACCCCGGCGAGCGGGGCCACGTG 3440
 A T H G Q V T Y T G S S Y G I S S T A A C T P A S A G H V 334
 TGGATGTCCAAGCAGCAGGCGCGCCGCGCAGCCSCCAGAGCCCCAGGCCCCCGGCCCCGAGGCGCCCGCCGCA 3520
 W M S K Q Q A P P P P P P Q Q P P Q A P P A P Q A P P Q 361
 GCCGAGGGCGGCCCCACAGCAGCCGGCGGCACCCCGCAGCAGCCACAGGCGCACACGCTGACCACGCTGAGCAGCG 3600
 P Q A A P P Q Q P A A P P Q Q P Q A H T L T T L S S 387
 AGCCGGGCCAGTCCCAGCGAACGCACATCAAGACGGAGCAGCTGAGCCCCAGCCACTACAGCGAGCAGCAGCACTCG 3680

CCCCAACAGATCGCCTACAGCCCCTTCAACCTCCCACACTACAGCCCCTCCTACCCGCCCATCACCCGCTCACAGTACGA 3760
 P Q Q I A Y S P F N L P H Y S P S Y P P I T R S Q Y D 441
 CTACACCGACCACCAGAACTCCAGCTCCTACTACAGCCACGGCGAGGCCAGGGCACCCGGCCTCTACTCCACCTTCACCT 3840
 Y T D H Q N S S S Y Y S H A A G Q G T G L Y S T F T 467
 ACATGAACCCCGCTCAGCGCCCCATGTACACCCCATCGCCGACACCTCTGGGGTCCCTTCCATCCCGCAGACCCACAGC 3920
 Y M N P A Q R P M Y T P I A D T S G V P S I P Q T H S 494
 CCCAGCACTGGGAACAACCCGTCTACACACAGCTCACTCGACCTTGAGGAGGCCTCCCACGAAGGGCGAAGATGGCCGA 4000
 P Q H W E Q P V Y T Q L T R P * 509
 GATGATCCCTAAAAATAACCGAAGAAAGAGAGACCAACCAGAATCCCTTTGGACATTTGTGTTTTTTTTGTTTTTTTATT 4080
 TTGTTTTGTTTTTCTTCTTCTTCTTCTTAAAGACATTTAAGCTAAAGGCAACTCGTACCCAAATTTCCAAGACAC 4160
 AAACATGACCTATCCAAGCGCATTACCCACTTGTGGCCAATCAGTGCCAGGCCAACCTTGGCTAAATGGAGCAGCGAAA 4240
 TCAACGAGAAACTGGACTTTTTAAACCCTCTTACAGAGCAAGCGTGGAGGATGATGGAGAATCGTGTGATCAGTGTCTAA 4320
 ATCTCTCGCCTGTTGGACTTTGTAATTATTTTTTAGCAGTAATTAAGAAAAAAGTCCCTCTGTGAGGAATATTTCTCT 4400
 ATTTTAAATATTTTATGATGTACTGTGTATGATTCAATACCATTTGAGGGGATTTATACATATTTTTTAGATAAAATTA 4480
 AATGCTCTTATTTTTCCAACAGCTAAACTACTCTTAGTTGAACAGTGTGCCCTAGCTTTTTCTTGCAACCAGAGTATTTTT 4560
 GTACAGATTTGCTTTCTCTTACAAAAAGAAAAAATCCTGTTGTATTAACATTTAAAAACAGAATTGTGTTATGTGA 4640
 TCAGTTTTGGGGTTAACTTTGCTTAATTCCTCAGGCTTTGCGATTTAAGGAGGAGCTGCCTTAAAAAATAAAGGCC 4720
 TTATTTTGCAATTTATGGGAGTAAACAATAGTCTAGAGAAGCAATTTGGTAAGCTTTATCATATATATATTTTTTAAAGAAG 4800
 AGAAAAACACCTTGAGCCTTAAAACGGTGTCTGCTGGGAAACATTTGCACTCTTTTAGTGCATTTCTCTCTGCTTTGCT 4880
 GTTCACTGCAGTCTTAAGAAAAGAGGTAAGCAAGGAGATGAAATCTGTTCTGGGAATGTTTTCAGCAGCCAAATA 4960
 AGTGCCCGAGCACACTGCCCCGGTTGCTTGCCTGGGCCCATGTGGAAAGGCAGATGCCTGCTCGCTTGTACCTGTGC 5040
 CTCTCAGAACCACAGCAGTTAACCTTCAAGACATTCACCTTGTCTAAAATTTATTTATTTTGTAAAGGAGAGGTTTTAATTA 5120
 AACAAAAAATTTCTTTTTTTTTTTTTTCCAATTTTACCTTCTTTAAAATAGGTTGTTGGAGCTTTCCTCAAAGGGT 5200
 ATGGTCATCTGTTGTTAAATATGTTCTTAACTGTAACCAGTTTTTTTTTATTTATCTCTTAAATCTTTTTTTATTATTA 5280
 AAAGCAAGTTTCTTTGTATTCCTCACCTAGATTTGTATAAATGCCTTTTTGTCCATCCCTTTTTCTTTGTTGTTTTG 5360
 TTGAAAACAACTGGAACTTGTTCCTTTTTTGTATAAATGAGAGATTGCAAATGTAGTGTATCACTGAGTCATTTGCA 5440
 GTGTTTTCTGCCACAGACCTTTGGGCTGCCTTATATTGTGTGTGTGTGTGGGTGTGTGTGTTTTGACACAAAAACAAT 5520
 GCAAGCATGTGTATCCATATTTCTCTGCATCTTCTTGGAGTGAGGGAGGCTACCTGGAGGGGATCAGCCCCTGACA 5600
 GACCTTAATCTTAATTACTGCTGTGGCTAGAGAGTTTGGAGATTGCTTTTTAAAAAAGACAGCAAACCTTTTTTTTTTATT 5680
 TAAAAAAGATATATTAACAGTTTGAAGTCAAGTAGAATAAATCTAAAGCACTCATAATATGGCCTCCTTCAATTTT 5760
 TGTATAAAGCAGATCTTTTTAAAAAGATACTCTGTAACCTAAGAAACCTGGCATTAAATCATATTTTGTCTTTAGT 5840
 AAAAGCTTTGGTTTTGTTGTTGTTGTTGTTTCACTTGTTCCTCCAGCCCAACCTTTTGTCTCTCCGTGAA 5920
 ACTTACCTTTCCCTTTTTCTTCTTTTTTTTTTGTATATTATTGTTTACAATAAATATACATTGCATTAATAAAGAA 6000
 AGTGGCCCTGTGGATTTATTCACCCAGTTCTCCTGTTGGATGATTTGGCAAATTTGAGCACAAAAGACATTGTGGAGTGC 6080
 TGATTGCTGTGATGTTTTTGTTTTTTATCAGCACCTTTGGCAATCCTAGAAAACAAGCTGAGGAAAAGACTGTATCTCCA 6160
 AAAATCTAGGCAGAAAATCTTGAAAAGTGCCTCCAATAGATCACACAGAAAATTACATGTCAGTAGTTGCTCACTCT 6240
 GGCAAAAATGTTTGTGTTGTTTTCATGACCTCATGCTTCAGGGCAAAGCTGTCCCTTGTGGAGGTCACAGGGAATATT 6320
 AGACCTGAAATCAGGAGCTTGTGAGTAGACAGAATGCCAGAGGGGTGACTTGTCTTATCTGGTTTTAATATGGCAAACCTT 6400
 CCGTTCCCTATGAACACATACCCAAGAAATGGAGTATCCACTTAAAAAGCAAAGGAAGCCAGAGAAAATCAGTGTCTACAG 6480
 GGAACCAGAGAGAAGCCTGTCGTATTAACCCATTAATGATTACAGAGCCTTCCAGATTTCTCTGTAGAGACAATGAAAGG 6560
 GGATGATTTTTCTGCTCCCTCAGTTTAACTC 6592



Séquence 2

GGAGAGCCGAAAGCGGAGCTCGAAACTGACTGGAAACTTCAGTGGCGCGGAGACTCGCCAGTTTCAACCCCGGAAACTTT 80
TCTTTGCAAGGAGGAGAAGAGAAGGGGTGCAAGCGCCCCACTTTTGTCTTTTTCTCCCTCCTCCTCCTCCTCAATTC 160
GCCTCCCCCACTTGGAGCGGGCAGCTGTGAAGTGGCCACCCCGCCTTCTAAGTGCTCGCCGCGGTAGCCGGCCGAC 240
GCGCCAGCTTCCCGGGAGCCGCTTGTCTCCGCATCCGGGCAGCCGAGGGGAGAGGCCCGCCCTCGAGTCCCGGAGCC 320
GCCGCGCTTCTCGCCTTTCCCGGCCACCAGCCCCCTGCCCGGGCCCGGTATGAATCTCCTGGACCCCTTCATGAAGA 400
M N L L D P F M K 10
TGACCGACGAGCAGGAGAAGGGCTGTCCGGCGCCCCAGCCCCACCATGTCCGAGGACTCCGCGGGCTCGCCCTGCCCG 480
M T D E Q E K G L S G A P S P T M S E D S A G S P C P 36
TCGGGCTCCGGTCCGACACCGAGAACACGGGCCCCAGGAGAACACGTTCCCCAAGGGCGAGCCCGATCTGAAGAAGGA 560
S G S G S D T E N T R P Q E N T F P K G E P D L K K E 63
GAGCGAGGAGACAAGTTCCCGTGTGCATCCGCGAGGCGGTGAGCCAGTGTCTCAAAGGCTACGACTGGACGCTGGTGC 640
S E E D K F P V C I R E A V S Q V L K G Y D W T L V 90
CCATGCCGGTGCAGCTCAACGGCTCCAGCAAGAACAAGCCGCACGTCAAGCGGCCCATGAACGCCTTCATGGTGTGGGCG 720
P M P V R V N G S S K N K P H V K R P M N A F M V W A 116
CAGGCGCGCGCAGGAAGCTCGCGGACCAGTACCCGCACCTGCACAACGCCGAGCTCAGCAAGACGCTGGGCAAGCTCTG 800
Q A A R R K L A D Q Y P H L H N A E L S K T L G K L W 143
GAGACTTCTGAACGAGAGCGAGAAGCGGCCCTTCGTGGAGGAGGCGGAGCGGCTGCGCGTGCAGCACAAGAAGGACCACC 880
R L L N E S E K R P F V E E A E R L R V Q H K K D H 170
CGGATTACAAGTACCAGCCGCGGCGGAGGAAGTCCGGTGAAGAACGGGCAGGCGGAGGAGGAGGCCACGGAGCAGACG 960
P D Y K Y Q P R R R K S V K N G Q A E A E A T E Q T 196
CACATYCCCCCAAGCCATCTTCAAGCGCTGCAGGCGGACTCGCCACACTCCTCCTCCGGCATGAGCGAGGTGCACTC 1040
H I S P N A I F K A L Q A D S P H S S S G M S E V H S 223
CCCCGGCGAGCACTCGGGCAATCCCAGGGCCACCACCCACCCACCCCAAAAACCGACGTGCAGCCGGGCAAGG 1120
P G E H S G Q S Q G P P T P P T T P K T D V Q P G K 250
CTGACCTGAAGCGAGAGGGGCGCCCTTGCCAGAGGGGGGCGAGACAGCCCCCTATCGACTTCCGCGAGCTGGACATCGGC 1200
A D L K R E G R P L P E G G R Q P P I D F R D V D I G 276
GAGCTGAGCAGCGACGTCACTCCAACATCGAGACCTTCGATGTCAACGAGTTTGACCAGTACCTGCCGCCAACGGCCA 1280
E L S S D V I S N I E T F D V N E F D Q Y L P P N G H 303
CCCCGGGGTCCCGGCCACGCACGGCCAGGTCACTACACGGGCAGCTACGGCATCAGCAGCACCCGCGGCCACCCCGGCGA 1360
P G V P A T H G Q V T Y T G S Y G I S S T A A T P A 330
GCGCGGVCACGTGTGGATGTTCAAGCAGCAGGCGCCGCCACCCCGCAGCAGCCCCACAGGCCCGCCGGCCCCCG 1440
S A G H V W M S K Q Q A P P P P P Q Q P Q A P P A P 356
CAGGCGCCCCCGCAGCCGAGCGCGGCCCCCCACAGCAGCCGGCGGCACCCCGCAGCAGCCACGGCGCACACGCTGAC 1520
Q A P P Q P Q A A P P Q Q P A A P P Q Q P Q A H T L T 383
CACGCTGAGCAGCGAGCCGGGCCAGTCCAGCGAACGCACATCAAGACGGAGCAGCTGAGCCCCAGCCACTACAGCGAGC 1600
T L S S E P G Q S Q R T H I K T E Q L S P S H Y S E 410
AGCAGCAGCACTCGCCCCAACAGATCGCCTACAGCCCCCTTCAACCTCCACACTACAGCCCCCTCTACCCGCCCATCACC 1680
Q Q Q H S P Q Q I A Y S P F N L P H Y S P S Y P P I T 436
CGCTCACAGTACGACTACACCGACCACCAGAACTCCAGCTCCTACTACAGCCACGCGGCAGGCCAGGGCACCGGCTCTA 1760
R S Q Y D Y T D H Q N S S S Y Y S H A A G Q G T G L Y 463
CTCCACCTTACCTACATGAACCCCGCTCAGCGCCCCATGTACACCCCATCGCCGACACCTCTGGGGTCCCTTCCATCC 1840
S T F T Y M N P A Q R P M Y T P I A D T S G V P S I 490
CGCAGACCCACAGCCCCAGCACTGGGAACAACCGTCTACACACAGCTACTGACCTTGAGGAGGCCCTCCACGAAGG 1920
P Q T H S P Q H W E Q P V Y T Q L T R P * 509
GCCAAGATGGCCGAGATGATCCTAAAAATAACCGAAGAAAGAGAGGACCAACCAGAATTCCCTTTGGACATTTGTGTTTT 2000
TTTGTTTTTTTATTTTGTGTTTTTCTTCTTCTTCTTCTTAAAGACATTTAAGCTAAAGGCAACTCGTACCCA 2080
AATTTCCAAGACACAAACATGACCTATCCAAGCGCATTACCCACTTGTGGCCAATCAGTGGCCAGGCCAACCTTGCTAA 2160
ATGGAGCAGCGAAATCAACGAGAACTGGACTTTTTAAACCCTCTTCAGAGCAAGCGTGGAGGATGATGGAGAATCGTGT 2240
GATCAGTGTGCTAAATCTCTGCTGCTTGGACTTTGTAATTTTTTTTTAGCAGTAATTAAGAAAAAAGTCCCTCTGT 2320
GAGGAATATCTCTATTTTAAATATTTTTAGTATGTACTGTGTATGATTCATTACCATTTTGAGGGGATTTATACATATT 2400
TTTAGATAAAATTAATGCTCTTATTTTTCCAACAGCTAAACTACTCTTAGTTGAACAGTGTGCCCTAGCTTTTCTTGCA 2480
ACCAGATATTTTTGTACAGATTTGCTTCTTACAAAAAGAAAAAATCCTGTTGTATTAACATTTAAAAACAGA 2560
ATTGTGTTATGTGATCAGTTTTGGGGTTAACTTTGCTTAATCTCAGGCTTTGCGATTTAAGGAGGAGCTGCCTTAA 2640
AAAAAATAAAGGCCTTATTTTGCATTTATGGGAGTAAACAATAGTCTAGAGAAGCATTGGTAAGCTTTATCATATATAT 2720
ATTTTTTAAAGAAGAGAAAAACACCTTGAGCCTTAAACCGGTGCTGCTGGGAAACATTTGCACTCTTTTACTGCATTTCC 2800
TCCTGCCTTTGCTTGTTCACCTGCAGTCTTAAAGAAAGAGGTAAGGCAAGCAAGGAGATGAAATCTGTTCTGGGAATGT 2880
TTCAGCAGCCAATAAGTGCCCGAGCACACTGCCCGCGTTGCCTGCCTGGGCCCATGTGGAAGGCAGATGCCTGCTCGC 2960
TCTGTACCTGTGCCTCTCAGAACACCAGCAGTTAACCTTCAAGACATTCCACTTGCTAAAATTTATTTATTTGTAAGGA 3040
GAGGTTTTAATTAACAAAAAATCTTTTTTTTTTTTTTTTCCAATTTTACCTTCTTTAAATAGGTTGTTGGAGC 3120
TTTCTCAAAGGTATGGTCACTGTGTTAAATATGTTCTTAACTGTAACCAGTTTTTTTTTTTATTTATCTTTAATC 3200
TTTTTTTTATTTAAAGCAAGTTTCTTTGTATTCTCACCTTAGATTTGTATAAATGCCTTTTTTGTCCATCCCTTTTTT 3280

CTGAGTCATTTGCAGTGTTCCTGCCACAGACCTTTGGGCTGCCTTATATTGTGTGTGTGTGTGGGTGTGTGTGTGTTC 3440
 GACACAAAAACAATGCAAGCATGTGTATCCATATTTCTCTGCATCTTCTCTTGGAGTGAGGGAGGCTACCTGGAGGGGA 3520
 TCAGCCCCTGACAGACCTTAATCTTAATTACTGCTGTGGCTAGAGAGTTTGAGGATTGCTTTTTAAAAAAGACAGCAA 3600
 CTTTTTTTTTTTATTTAAAAAAGATATATTAACAGTTTTAGAAAGTCAGTAGAATAAAATCTTAAAGCACTCATAATATGG 3680
 CATCCTTCAATTTCTGTATAAAAAGCAGATCTTTTTAAAAAGATACTTCTGTAACCTAAGAAACCTGGCATTAAATCATA 3760
 TTTTGTCTTTAGGTAAAAGCTTTGGTTTGTGTTTCGTGTTTTGTTTGTTCACCTTGTTCCTCCCAGCCCCAAACCTTTT 3840
 GTTCTCTCCGTGAAACTTACCTTTCCCTTTTTCTTTCTTTTTTTTTTTTTTGTATATTATTGTTTACAATAAATATACAT 3920
 TGCATTAAAAAG 3932

Séquence 3

TTTTTGTGTCCTCTCGTTTTGTGACATAAAAGGTCAATGAAAAAATTGGCGATTAAGTCAAATTCGCATTTTTCAGGACA 80
 GCAGTAGAGCAGTCAGGGAGGCAGATCAGCAGGGCAAGTAGTCAACGTTACTGAATTACCATGTTTTGCTTGAGAATGAA 160
 TACATTGTCAAGGCTACTAGGGGGTAGGCTGGTTGGGCGGGTTGAGGGGGTGTGAGGGCGGAGAAATGCAAGTTTCATT 240
 ACAAAAAGTTAACGTAACAAAGAATCTGGTAGAAGTGAGTTTTGGATAGTAAAATAAGTTTCGAACCTCTGGCACCTTTCAA 320
 TTTTGTGCGCACTCTCCTTGTTTTTGACAATGCAATCATATGCTTCTGCTATGTTAAGCGTATTC AACAGCGATGATTACA 400
 M Q S Y A S A M L S V F N S D D Y 18

GTCCAGCTGTGCAAGAGAATATTCCCCTCTCCGGAGAAGCTCTTCCCTTCTTTGCACTGAAAGCTGTAACCTTAAGTAT 480
 S P A V Q E N I P A L R R S S S F L C T E S C N S K Y 44

CAGTGTGAAACGGGAGAAAACAGTAAAGGCAACGTCCAGGATAGAGTGAAGCGACCCATGAACGCATTCATCGTGTGGTC 560
 Q C E T G E N S K G N V Q D R V K R P M N A F I V W S 71
 TCGCGATCAGAGGGCGAAGATGGCTCTAGAGAATCCCAGAATGCGAAACTCAGAGATCAGCAAGCAGCTGGGATACCAGT 640
 R D Q R R K M A L E N P R M R N S E I S K Q L G Y Q 98
 GGAAAATGCTTACTGAAGCCGAAAAATGGCCATTCTTCCAGGAGGCACAGAAATTACAGGCCATGCACAGAGAGAAATAC 720
 W K M L T E A E K W P F F Q E A Q K L Q A M H R E K Y 124
 CCGAATTATAAGTATCGACCTCGTCGGAAGGCGAAGATGCTGCCGAAGAATTGCAGTTTGCTTCCCGCAGATCCCGCTTC 800
 P N Y K Y R P R R K A K M L P K N C S L L P A D P A S 151
 GGTACTCTGCAGCGAAGTGCAACTGGACAACAGGTTGTACAGGGATGACTGTACGAAAGCCACACACTCAAGAATGGAGC 880
 V L C S E V Q L D N R L Y R D D C T K A T H S R M E 178
 ACCAGCTAGGCCACTTACCGCCCATCAACGCAGCCAGCTCACCGCAGCAACGGGACCGCTACAGCCACTGGACAAAGCTG 960
 H Q L G H L P P I N A A S S P Q Q R D R Y S H W T K L 204
 TAGGACAATCGGGTAACATTGGCTACAAAGACCTACCTAGATGCTCCTTTTTTACGATAACTTACAGCCCTCACTTTCTTA 1040
 *

TGTTTAGTTTCAATATTGTTTTCTTTTCTCTGGCTAATAAAGGCCTTATTCATTTTCAGTTTTACTGGTATTTTCATTTTAA 1120
 ACTTAAATTCAGACAAGTTGTGTCAACACGATTAACATGCAAAGAAATAAGACATCCAGAAGTGAGCCTGCCTATGTTT 1200
 GTGGCCGTGAGTACTAACTTGATACAAACGGACACTGTGGCTTACTTTAAATGCTCTAATGAGAAACACACTTGAAAA 1280
 TTGTACCAAAAAAATCACACTTCTATATGCAGCGTGTTA 1320