

SUGGESTIONS PÉDAGOGIQUES : CLASSE TERMINALE

SÉRIE S

INTRODUCTION : LA BANQUE DE DONNÉES ET LES NOTIONS DU PROGRAMME

Les séquences et documents réunis peuvent être utilisés pour construire les notions des diverses parties du programme de biologie de terminale S. Certaines de ces séquences sont également fournies dans le logiciel *Phylogène* et sont susceptibles d'être traitées à l'aide de fonctions propres à ce logiciel, comme celle de construction d'arbre par exemple. Les deux logiciels sont donc complémentaires.

1 - Pour la partie « **Parenté entre êtres vivants actuels et fossiles** », le choix des molécules permet de rechercher les relations de parenté au sein d'ensembles plus ou moins étendus. Ainsi, le gène **CDC2** qui code pour un polypeptide jouant un rôle majeur dans la réalisation du cycle cellulaire peut être utilisé pour les **relations de parenté au sein des eucaryotes**. Les **gènes homéotiques** permettent de rechercher les parentés chez les **métazoaires bilatéraux**. Les gènes codant pour les **globines alpha et bêta** se prêtent à l'établissement de **phylogénies chez les Vertébrés** et le gène codant pour le **pigment visuel** de courte longueur d'onde (opsine S ou « bleue ») peut être utilisé pour préciser les relations de parenté chez les **Primates**. Enfin, la banque contient des séquences de **l'ADN mitochondrial** de Primates actuels et de l'Homme de Neandertal. Ces séquences peuvent servir pour préciser les relations de parenté chez les **Primates** mais surtout avec l'exploitation de l'ADN fossile des neandertaliens permettent de discuter des relations entre **l'Homme de Neandertal et Homo sapiens**.

2 - En ce qui concerne la partie « **Stabilité et variabilité des génomes et évolution** », les séquences et documents fournis permettent d'envisager les différents types d'innovations génétiques et les mécanismes susceptibles d'assurer leur maintien dans les populations.

- L'exemple des **globines**, classique mais très documenté, est le plus riche et peut être utilisé pour établir les principales notions. En effet, l'analyse des divers allèles du gène bêta conduit à faire le point sur les différents types de **mutations ponctuelles** et celle des gènes codant pour les différentes chaînes de globine permet d'établir la notion de **famille multigénique** et donc celle de production de nouveaux gènes par duplication. La comparaison des séquences du gène alpha ou bêta chez différentes espèces ou celle des différents gènes de la famille multigénique dans l'espèce humaine mettent en évidence des sites conservés et des sites à forte variabilité. L'interprétation de ces sites à l'aide des documents fournis conduit aux idées de **conservation de mutations neutres** (par dérive génique) et **d'élimination des mutations affectant la fonction des globines**. Enfin, l'exemple classique de l'avantage procuré par la possession de l'allèle HbS chez les hétérozygotes en région paludéenne, permet de comprendre la **permanence et l'extension d'un allèle par sélection naturelle** dans certaines conditions d'environnement.
- Les exemples des gènes codant pour **l'alpha-antitrypsine** et la **G6PD** permettent aussi d'établir les différents types de **mutations ponctuelles** et surtout d'établir des **filiations entre allèles** et donc de reconstruire au moins en partie l'histoire d'un gène au sein des populations humaines. En outre, grâce à la riche documentation associée, l'exemple de la G6PD renforce celui de l'hémoglobine S pour faire saisir comment dans certaines conditions d'environnement, certains allèles se répandent par **sélection positive** dans certaines populations.
- L'importance du processus de **duplication suivi de mutations ponctuelles** des gènes dupliqués peut être appréhendée à partir de divers exemples de familles multigéniques : celle des **opsines** chez les Primates (à l'origine de la possibilité de vision des couleurs), celle des hormones hypophysaires LH, FSH, TSH et placentaire HCG, celle des hormones hypophysaires Gh (hormone de croissance), HPRL (prolactine) et HLP (hormone lactogène placentaire).
- Enfin, un des exemples les plus remarquables associant innovations génétiques et sélection naturelle est celui de **l'acquisition de la résistance aux insecticides par les Moustiques**. Sans traiter complètement le problème, on y voit comment des duplications de gènes codant pour les estérases, enzymes capables de dégrader les insecticides organophosphorés, arrivent à se répandre dans les régions où l'épandage de ces insecticides a été important de sorte qu'une grande partie de la population de Moustiques est devenue résistante.

3 - Pour la partie « **Procréation** », les documents fournis sont relatifs à **des cas cliniques d'infertilité** due à des allèles mutés de gènes intervenant soit dans la différenciation sexuelle au cours de la vie embryonnaire ou fœtale, soit dans la commande hormonale de l'appareil génital à la puberté. Les données morphologiques, anatomiques et les tests biologiques doivent permettre de faire des hypothèses sur le gène en cause dans les différents cas cliniques. Les séquences des allèles des différents gènes présents chez l'individu comparées à celles d'allèles fonctionnels de référence permettent de tester les hypothèses émises.

4 - Pour la partie « **Immunologie** », les séquences fournies sont celles des chaînes lourdes et légères **d'immunoglobulines anti VIH**. La comparaison des séquences de ces chaînes permet de dégager la notion de partie constante et de partie variable de chaque type de chaîne et donc de faire l'hypothèse que la spécificité des anticorps est liée aux régions variables (et plus précisément aux zones hypervariables). Cela peut être testé à l'aide du logiciel *Rasmol* ou *RasTop* qui permet de localiser sur la structure spatiale d'une immunoglobuline les sites de fixation antigénique et de constater qu'ils correspondent aux régions hypervariables des chaînes d'immunoglobulines. Un travail du même type peut être conduit à partir des séquences des récepteurs des lymphocytes T.

5 - En ce qui concerne **l'Enseignement de spécialité**, la banque de données fournit des séquences relatives à une famille où des individus sont **albinos**. L'exemple est relativement complexe dans la mesure où l'albinisme dans l'arbre généalogique fourni a pour origine deux gènes différents, mais il est possible d'avoir une démarche très progressive. Surtout, les données fournies permettent de saisir comment agissent les enzymes de restriction et comment les fragments d'ADN obtenus par l'action de ces enzymes peuvent être séparés par électrophorèse.

PARENTÉ ENTRE ÊTRES VIVANTS ACTUELS ET FOSSILES – PHYLOGENÈSE, ÉVOLUTION

Relations de parenté au sein du vivant – Gène CDC2

Informations scientifiques

Des informations complémentaires sont disponibles sur le site bmedia :

<http://www.snv.jussieu.fr/bmedia/cyclecellBM/index.htm>

Le cycle cellulaire

Un cycle cellulaire constitué d'une interphase et d'une mitose est observable chez tous les eucaryotes. L'interphase comprend elle-même trois phases : la phase G1 (ou Gap 1), la phase S (réplication de l'ADN) et la phase G2 (ou Gap 2). La durée totale d'un cycle cellulaire varie beaucoup d'une cellule à l'autre. La durée de chacune des phases du cycle cellulaire peut être calculée si l'on connaît le pourcentage de cellules en phases S (incorporation de BrdU), G1 et G2 (estimation de la quantité moyenne d'ADN par noyau) et la durée moyenne du cycle cellulaire (elle peut être calculée par une expérience d'incorporation de thymidine tritiée suivie d'une autoradiographie : la durée du cycle cellulaire étant égale au temps nécessaire pour obtenir tous les noyaux marqués). Des mesures effectuées chez différents organismes montrent que, d'une façon générale, la phase G1 a la durée la plus variable du cycle cellulaire, et cette variabilité est responsable des différences de durée des cycles cellulaires observés.

La mise en évidence d'un contrôle du cycle cellulaire

Cette mise en évidence s'est faite tout d'abord chez les Levures (*Saccharomyces cerevisiae*, Levure bourgeonnante, et *Schizosaccharomyces pombe*, Levure fissipare). L'étude des mutants présentant des perturbations du cycle cellulaire a permis la découverte de gènes codant pour des protéines impliquées dans le contrôle du cycle cellulaire. Ces protéines étant très conservées au cours de l'évolution, on les a ensuite recherchées dans les cellules de Mammifères et d'autres Vertébrés, ce qui a permis d'identifier les gènes impliqués dans le contrôle du cycle cellulaire chez l'Homme, la Souris, le Xénope, etc.

Les mutations étudiées chez les Levures

S. cerevisiae et *S. pombe* ont des cycles cellulaires similaires. Leur reproduction possible à l'état haploïde a facilité l'étude des mutations. D'autre part, pour maintenir une taille moyenne constante, la durée du cycle cellulaire doit évaluer la durée de croissance nécessaire au doublement de la taille de la cellule. Or, la croissance cellulaire dépend beaucoup des nutriments disponibles dans l'environnement ; il existe donc un contrôle du cycle cellulaire en fonction de la taille cellulaire chez les Levures : la cellule ne pourra se diviser que si elle a acquis une taille suffisante. Des mutations dans les gènes responsables de ce contrôle entraînent l'existence de cellules de taille anormale, trop grandes ou trop petites.

Les mutations de deux types de gènes ont été étudiées :

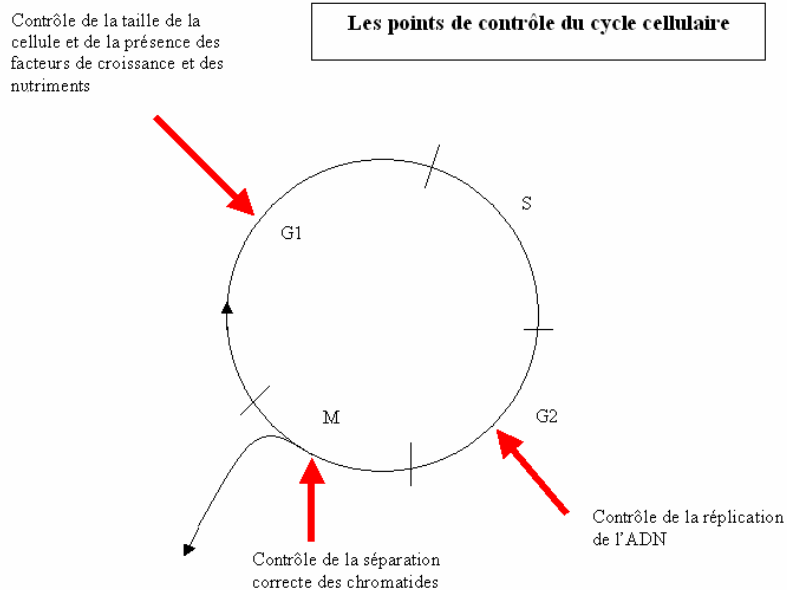
- celles des gènes *cdc* (Cell division cycle) qui codent pour des protéines kinases qui, en association avec des cyclines spécifiques, déclenchent le passage d'un stade du cycle cellulaire au suivant (par exemple, le passage du stade G2 au stade M pour la protéine kinase codée par le gène *cdc2*). Les mutations de ces gènes qui ont été le plus utiles pour décrypter le cycle cellulaire sont des mutations conduisant à des Levures dont le phénotype est conditionnel : la protéine « mutante » est fonctionnelle uniquement dans des conditions spécifiques. La plupart des mutants sont thermosensibles : la protéine mutante n'est pas fonctionnelle à température élevée mais permet le déroulement du cycle cellulaire à basse température. Une souche cellulaire de mutant *cdc2* thermosensible peut donc pousser uniquement à basse température, condition permissive ;
- celles des gènes qui codent pour des protéines qui agissent sur les protéines codées par les gènes *cdc* soit en les inhibant, soit en les activant (par phosphorylation ou déphosphorylation). L'expression de ces gènes dépend de facteurs divers, notamment de la taille de la cellule. Par exemple, le gène *wee* code pour une protéine qui inhibe la protéine kinase codée par le gène *cdc2*. Tant que cette protéine inhibitrice est produite, le passage du stade G2 au stade M ne peut avoir lieu. Lorsque la taille de la Levure est suffisante, le gène *wee* cesse de s'exprimer et l'entrée en mitose peut avoir lieu. Un mutant *wee* déficient ne bloque pas la protéine kinase codée par *cdc2* de sorte que la levure entre en mitose de façon prématurée (Levures fissipares de petite taille ; voir le thème de seconde).

C'est chez la Levure fissipare (*S. pombe*) que le gène *cdc2* a été découvert. Une mutation de ce gène empêche la mitose : il provoque un arrêt du cycle en empêchant l'entrée en phase G1 ou le passage G2/M (l'activité même du gène *cdc2* est régulée par les produits de l'expression d'autres gènes, notamment les gènes *wee1*, *cdc25* et *cdc1*). Chez la Levure bourgeonnante, le gène *cdc28* est équivalent au gène *cdc2* de la Levure fissipare (on regroupe maintenant les deux protéines produites par chacun de ces gènes *cdc2* et *cdc28* sous le terme *p34cdc2*).

Les points de contrôle du cycle cellulaire

Il existe trois points de contrôle principaux :

- au début de la phase G1 : les facteurs de croissance et les nutriments doivent être présents, la taille de la cellule doit être suffisante, pour que le cycle se poursuive ;
- au niveau du passage S/G2 : les cellules stoppent leur cycle cellulaire si l'ADN est incomplètement répliqué ou endommagé et le reprennent après réparation ;
- au cours de la mitose, à la fin de la métaphase : les cellules vérifient que les chromosomes sont correctement attachés aux fibres du fuseau avant que ne démarre la séparation des chromatides.



(Si des dommages dans l'ADN ou les chromosomes sont constatés et qu'ils ne sont pas réparables, la cellule s'engage sur la voie de l'apoptose ou mort cellulaire programmée.)

Ces points de contrôle agissent sur la poursuite ou non du cycle cellulaire par l'intermédiaire de protéines kinases cyclines dépendantes. Celles-ci résultent de l'association de protéines kinases (*cdc* ou *cdk*) avec des cyclines. Il existe de très nombreuses cyclines, et leur taux varie beaucoup au cours du cycle comme leur nom l'indique ; ce taux résulte du bilan entre leur synthèse et leur dégradation. Le taux de protéines kinases *cdk* (ou *cdc*) est au contraire relativement constant.

La succession des événements qui aboutit à la division cellulaire est donc en grande partie déterminée par l'expression séquentielle des cyclines ; elles agissent pour initier les phases du cycle cellulaire et aussi pour terminer ces phases.

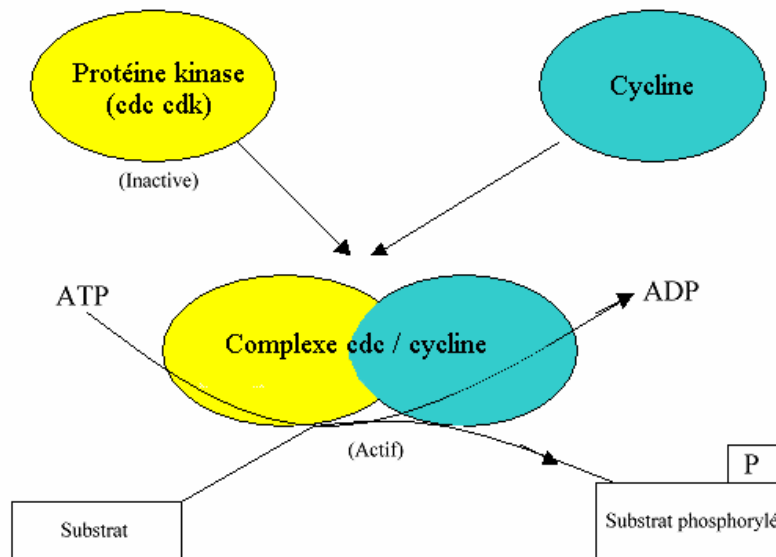
L'activité des protéines kinases cyclines dépendantes est ainsi réglée par le niveau d'expression des cyclines, mais aussi par l'action de protéines les déphosphorylant ou les phosphorylant et par des protéines inhibitrices, les CKI.

Remarques

Les cyclines sont dégradées à la fin de la phase dans laquelle elles interviennent, ce qui évite à la cellule de reprendre la même phase et impose une progression du cycle cellulaire.

Une carence en nutriments réduit la synthèse des cyclines *cln* par rapport à leur taux de dégradation, ce qui réduit leur concentration qui devient alors inférieure au seuil nécessaire pour entraîner l'activité de la protéine kinase *cdc2* (ou *p34cdc2*). C'est ainsi que l'environnement peut intervenir sur la durée du cycle cellulaire chez la Levure *S. cerevisiae*.

Mode d'action des protéines kinases cycline dépendante (cdc ou cdk)



Les cdk sont donc des enzymes qui catalysent la phosphorylation de protéines cibles jouant un rôle dans les événements du cycle cellulaire : fragmentation de l'enveloppe nucléaire, compaction des chromosomes, etc. De cette phosphorylation résulte un changement de conformation de ces protéines cibles qui entraîne des propriétés nouvelles.

Chez les Levures

Les études réalisées chez la Levure bourgeonnante (*S. cerevisiae*) montrent que selon les phases du cycle, ce n'est pas la même cycline qui sert d'activateur de la protéine kinase cdc2 (ou cdk1) :

- cyclines cln 1, cln 2 cln 3 pour pouvoir démarrer la phase G1 (complexe cdc2/cln1, cdc2/cln2, cdc2/cln3) ;
- cyclines clb 1, clb 2, clb 3, clb 4 pour finir la mitose (complexe cdc2/clb1, cdc2/clb2, cdc2/clb3, cdc2/clb4) ;

Chez la Levure *S. pombe*, elle est impliquée dans le démarrage du cycle en G1 et dans la transition G2-M.

Chez la Drosophile, le Xénope et les Mammifères

Plusieurs associations cycline/protéine kinase contrôlent le cycle cellulaire :

- cdk4/cycline D et cdk2/cycline E pour pouvoir démarrer la phase G1 ;
- cdk2/cycline A pour entrer en phase S ;
- cdk1/cycline B pour entrer en mitose. L'activité du complexe cdk1/cycline B permet la condensation de l'ADN, le désassemblage de l'enveloppe nucléaire et le réarrangement du cytosquelette. En fin de mitose, l'inactivation de ce complexe par d'autres protéines kinases (*wee1*, *myt 1*) permet la fin de la mitose.

Le gène cdc2 (ou cdk1)

Le gène *cdc2* (*cell division cycle*), chez la Levure, ou *cdk1* (*cyclin dependant kinase*), chez les Vertébrés, code pour une protéine kinase cycline dépendante, c'est-à-dire qui a besoin, pour être activée, de se lier à une cycline.

Il est présent chez tous les Eucaryotes et il a été isolé et séquencé chez de nombreux organismes unicellulaires ou pluricellulaires. L'homologie est très importante entre tous ces gènes.

D'autre part, des expériences de transgénèse ont montré que le gène *cdc2* d'une espèce, transféré dans une cellule d'une autre espèce, peut y régir le cycle cellulaire : ainsi, des Levures thermosensibles de la souche *Schizosaccharomyces Pombe* dans lesquelles on a transféré le gène *cdc2* humain deviennent capables de se diviser à des températures élevées.

Ces observations et expériences sont en faveur d'une origine commune et ancienne des mécanismes de contrôle du cycle cellulaire et mettent en évidence la grande unité du monde vivant.

Les travaux sur le gène *CDC2* et sur les autres gènes régulateurs du cycle cellulaire ont valu à P. Nurse, L. Harwell et T. Hunt le prix Nobel de médecine 2001.

Pistes d'exploitation pédagogique des données fournies

Les données moléculaires peuvent être utilisées pour établir des relations de parenté entre les êtres vivants. Ces données s'avèrent bien utiles, notamment quand les autres types de données (anatomiques, morphologiques ou embryologiques) ne sont pas utilisables ; c'est le cas notamment lorsque l'on veut préciser des relations de parenté entre des organismes très différents.

Le gène CDC2 est présent chez tous les organismes eucaryotes, et l'utilisation de données relatives à cet exemple sensibilise à l'idée que les mécanismes fondamentaux de la vie cellulaire sont partagés par tous les organismes eucaryotes.

Séquences et documents

Fichiers des séquences

Dans la banque de thèmes d'étude, le chemin **Parenté entre êtres vivants actuels et fossiles - Phylogénèse, évolution/Relations de parenté au sein du vivant/Gène CDC2** permet d'atteindre :

- **Gène CDC2 chez différentes espèces** qui charge le fichier genescdc2.edi affichant les séquences nucléiques strictement codantes du gène CDC2 chez quelques êtres vivants : Levure, Arabette, Chou, Blé, Maïs, Drosophile, Homme, Xénope, Rat, Souris, Poisson rouge, Grenouille, Poulet, Étoile de mer, Oursin ;
- **Protéine CDC2 chez différentes espèces** qui charge le fichier protCDC2.edi affichant les séquences protéiques du CDC2 correspondant aux séquences nucléiques.

Documents fournis

Dans la banque de documents, le chemin **Parenté entre êtres vivants actuels et fossiles - Phylogénèse, évolution/Relations de parenté au sein du vivant** permet d'atteindre **Gène CDC2** pour charger le fichier cdc2.bmp affichant un texte présentant l'importance du gène CDC2 et précisant son rôle. Ce document apporte un argument supplémentaire à l'unité du vivant en évoquant des expériences de transgénèse réussies.

La présence du gène CDC2 chez tous ces êtres vivants traduit une origine commune. La comparaison deux à deux des séquences nucléiques d'une part et protéiques d'autre part permet d'établir le tableau présenté à la page suivante.

	Arabette	Blé	Chou	Drosophile	Étoile de mer	Grenouille	Homme	Levure	Maïs	Oursin	Poisson rouge	Poulet	Rat	Xénope
Arabette	100													
Blé	82,7	100												
Chou	96,3	82	100											
Drosophile	62	60,9	63	100										
Étoile de mer	63,3	62,7	63,3	69	100									
Grenouille	58,3	57	57,9	62,9	65,2	100								
Homme	65,7	63,6	66	71,7	74,1	78,8	100							
Levure	62,6	62,6	63	60,6	63	58,9	63,6	100						
Maïs	81,6	92,9	81	61,6	63,3	59,5	64,6	63,3	100					
Oursin	60,5	60,1	61,1	68,4	75,1	64,1	72,1	59,8	60,1	100				
Poisson rouge	63,9	63,6	64,6	69,2	74,8	77,2	82,8	61,9	63,9	70,2	100			
Poulet	65,3	62	64,4	69	71,6	77,9	91,4	62	62,4	70	82,8	100		
Rat	65,3	64	65,7	71,7	73,7	79,1	97,6	64	63,6	72,7	83,8	91,6	100	
Xénope	65,2	62,9	64,6	68,9	71,5	87,7	86,8	61,9	63,6	71,2	84,8	87,4	85,4	100

Matrice des identités (en %) entre les protéines cdc2 fournies (obtenue à partir d'un alignement avec discontinuité)

NB : Ces résultats ont été obtenus en choisissant une séquence de référence (la première) et en maintenant toutes les autres dans l'ordre alphabétique des noms. Du fait de l'algorithme de comparaison, ces résultats peuvent être légèrement différents de ceux obtenus par comparaison des séquences deux à deux.

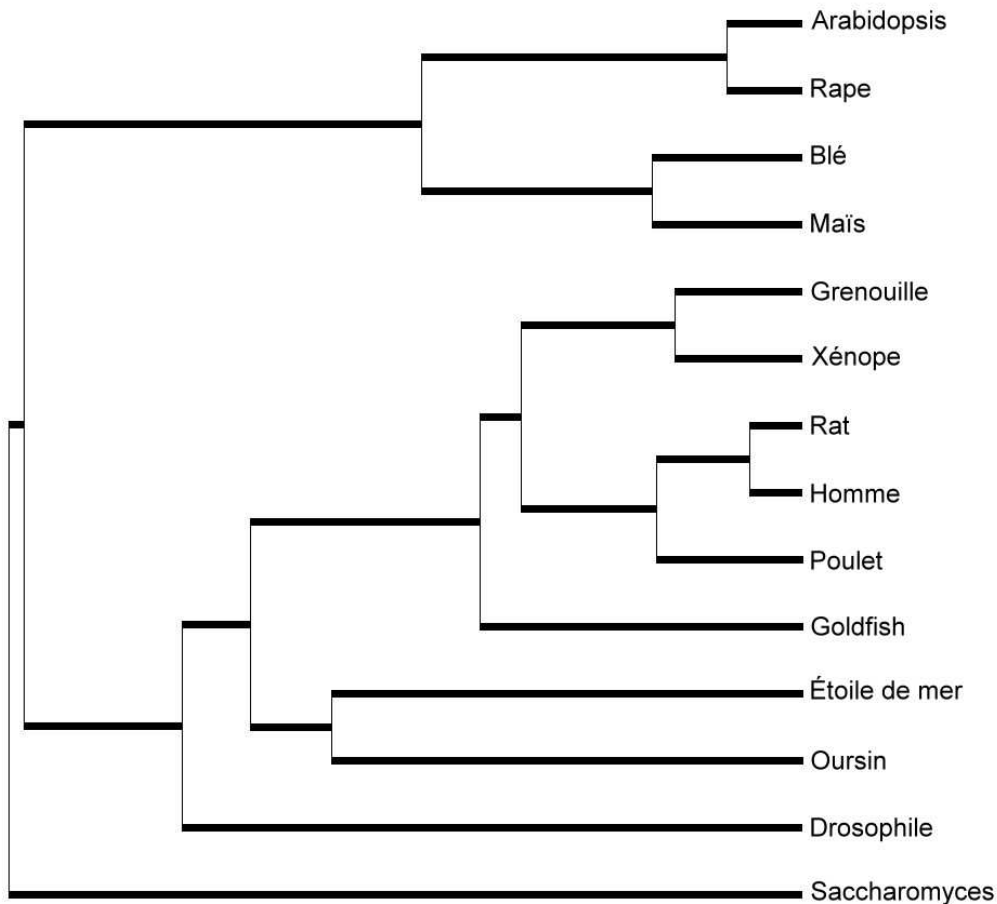
Remarques

- D'un point de vue pratique :
pour obtenir le degré d'identité entre les séquences comparées, il faut cliquer dans la première case de la ligne traitement, de façon à y faire apparaître la petite flèche rouge. Il faut ensuite cliquer sur l'icône avec un « I » qui donne l'information sur la ligne pointée. Le degré d'identité est exprimé en pourcentage.
- D'un point de vue conceptuel :
pour établir une phylogénie à partir de données moléculaires, on commence toujours par relier les éléments qui sont le plus proches, qui présentent le moins de différences, car ce sont eux qui ont l'ancêtre commun le plus récent.

La prise en compte du degré d'identité permet ensuite de préciser les parentés entre ces molécules, donc entre les organismes qui les possèdent. On applique alors une méthode phénétique, basée sur le seul nombre d'identités, sans tenir compte d'une quelconque polarité des états. Le raisonnement est le suivant : les différences constatées sont dues à des mutations ; or ces mutations s'accumulent au cours du temps et on peut donc considérer que plus il y a de différences, et plus de temps s'est écoulé depuis le dernier ancêtre commun,

et inversement ; donc, on peut finalement dire que plus le degré d'identité est important et plus la parenté est grande.

À partir de ces matrices, on peut construire un arbre phylogénétique semblable au suivant obtenu avec le logiciel *Phylogène* (représentation par la méthode UPGMA) :



La construction manuelle d'un tel arbre serait longue et fastidieuse. On peut se limiter à un choix d'espèces en fonction de problèmes à résoudre :

- rechercher qui des Insectes ou des Échinodermes sont les plus proches parents des Vertébrés ;
- établir les relations de parenté entre les quatre plantes à fleurs et voir si elles confortent le rangement de ces plantes dans deux groupes, les Crucifères et les Poacées (Graminées), groupes établis à partir des caractéristiques des Fleurs notamment ;
- rechercher si la classification des Vertébrés à partir de ces données moléculaires est conforme à celle obtenue à partir de données morphologiques et anatomiques.

Relations de parenté au sein du vivant – Gènes homéotiques

Informations scientifiques

Cf. page 45 (classe de seconde)

Pistes d'exploitation pédagogique des données fournies

L'exploitation des données sur les homéoboîtes et les homéodomains des gènes homéotiques permet d'aborder la notion d'unité du vivant et d'établir des relations de parenté entre certains organismes. Ainsi, des parentés pourront être établies entre des organismes très différents tels un Vertébré, un Arthropode, un Crustacé ou un Ver parasite, alors que les données anatomiques et morphologiques étaient difficilement utilisables.

Séquences et documents

Fichiers des séquences

Dans la banque de thèmes d'étude, le chemin **Parenté entre êtres vivants actuels et fossiles – Phylogénèse, évolution/Relations de parenté au sein du vivant/Gènes homéotiques** pour accéder à :

- **Homéobox (nucléique)** afin de charger les divers fichiers Bx-box.edi affichant les séquences nucléiques des homéoboîtes de quelques gènes homéotiques de Drosophile (Antp, Dfd, Lab, Scr, Pb, Ubx, AbdA, AbdB), d'Homme (Hox B4, B6, B7, B9, C6), de Souris (Hox B4, B6, B7, B9, C6), de Xénope (Hox B4, B7) ;

- **Homéodomaine (protéique)** afin de charger les divers fichiers Bx-domaines.edi affichant les séquences des domaines protéiques de quelques gènes homéotiques de Drosophile (gènes AbdA, AbdB, Antp, Lab, Scr, Dfd, Pb et Ubx), de Souris (Hox B4, B6, C6, B7, B9), d'Homme (Hox B4, B6, C6, B7, B9), de Xénope (Hox B4, B7, B9), de Sacculine (AbdB, Antp) et de Dicyemie (Antp).

La Souris et l'Homme sont des Vertébrés mammifères, le Xénope est un Vertébré amphibien, la Drosophile est un Arthropode Insecte, la Sacculine est un Arthropode crustacé, le Dicyemie est un animal parasite des Calmars et des Pieuvres, très spécialisé, Métazoaire du groupe des Lophotrochozoaires (ce groupe inclut les Annélides, les Brachiopodes, les Mollusques et les Bryozoaires).

Documents fournis

Dans la banque de documents, le chemin **Parenté entre êtres vivants actuels et fossiles – Phylogénèse, évolution/Relations de parenté au sein du vivant/Gènes homéotiques** permet d'atteindre :

- **Mutants homéotiques de Drosophile** qui charge le fichier phenotypeshomeo.jpg présentant la description des phénotypes de mutants homéotiques permet de comprendre l'action des gènes homéotiques. Ces mutations entraînent la formation d'un organe parfaitement constitué, mais pas au bon endroit. Les gènes homéotiques sont donc des gènes qui vont intervenir au cours du développement embryonnaire pour permettre la mise en place des différents organes. Ils sont responsables de l'édification du plan d'organisation de l'organisme ;

- **Localisation chromosomique** qui charge le fichier chromosomeshomeo.jpg affichant la disposition des gènes sur les chromosomes et permettant de faire plusieurs constats : les gènes sont organisés en complexes (un complexe de gènes disposés en deux groupes appelés parfois complexes antennapedia et bithorax, situés sur le même chromosome chez la Drosophile, et quatre complexes situés sur des chromosomes différents chez la Souris), la disposition des gènes sur les chromosomes correspond à la disposition des régions d'expression de ces gènes dans l'organisme, il y a une correspondance qui a pu être établie entre les gènes de la Drosophile et les gènes de la Souris (ces correspondances sont indiquées par les couleurs) ;

- **Expérience de transgénèse** qui charge le fichier transgenese.bmp : ces expériences de transgénèse réussies sont un argument supplémentaire en faveur de l'homologie des gènes homéotiques de Vertébrés et d'Arthropodes, et donc de l'unité du vivant ;

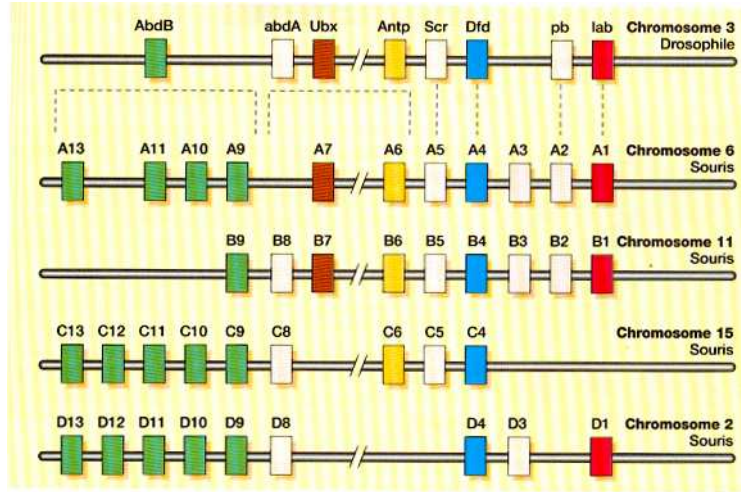
- **Complexe Homeodomaine/ADN** qui charge le fichier complexeADNHomeo.jpg illustrant le mode d'action des gènes homéotiques. La figure permet de comprendre comment agissent les gènes homéotiques : ils codent pour une protéine dont une partie, appelée homéodomaine, est capable de se fixer sur l'ADN, permettant ainsi la régulation de l'expression de certains gènes.

On se limite ici aux séquences nucléiques des homéoboîtes et des séquences protéiques des homéodomains car leur similitude est très grande d'un gène homéotique à l'autre, alors que cette homologie est impossible à mettre en évidence si l'on compare les séquences codantes entières de ces gènes (ou les protéines entières). En fait, les gènes homéotiques ne se ressemblent qu'au niveau des homéoboîtes codant pour les homéodomains protéiques. D'une façon générale, les conclusions obtenues seront les mêmes, que l'on travaille sur la comparaison des séquences protéiques des homéodomains ou sur la comparaison des séquences nucléiques des homéoboîtes.

Cette similitude de séquences suggère une origine commune, c'est-à-dire une séquence ancestrale commune, donc une origine commune pour les êtres vivants qui les possèdent, c'est-à-dire l'Homme, la Souris, le Xénope, la Drosophile, la Sacculine et le Dicyemie.

Au-delà de l'idée d'origine commune des animaux, cette similitude permet d'imaginer une apparition très précoce des gènes gouvernant la mise en place du plan d'organisation.

On compare les gènes qui sont « équivalents » chez la Drosophile et les Vertébrés. Ces « équivalences » peuvent être repérées sur le document présentant la disposition des gènes homéotiques chez la Drosophile et chez la Souris. Ainsi, on pourra comparer le gène Dfd de la Drosophile avec les gènes *hoxA4*, *B4*, *C4* et *D4* des Vertébrés, le gène *Ubx* de la Drosophile avec les gènes *A7* et *B7* le gène *Antp* avec les gènes *A6*, *B6* et *C6*.



Comparaison des domaines protéiques

Résultat des comparaisons des domaines protéiques codés par les gènes Dfd de Drosophile et Hox-B4 de trois Vertébrés : on obtient 86,7 % d'identité globale.

Le tableau ci-dessous précise les identités entre les séquences prises deux à deux :

	DP-B4 Homme	DP-B4 Souris	DP-B4 Xénope	DP-Dfd Drosophile
DP-B4 Homme	100 %			
DP-B4 Souris	100 %	100 %		
DP-B4 Xénope	95 %	95 %	100 %	
DP-Dfd Drosophile	88,3 %	88,3 %	88,3 %	100 %

Résultat des comparaisons des domaines protéiques codés par les gènes Ubx de Drosophile et Hox-B7 de trois Vertébrés : on obtient 88,3 % d'identité globale.

Le tableau ci-dessous précise les identités entre les séquences prises deux à deux :

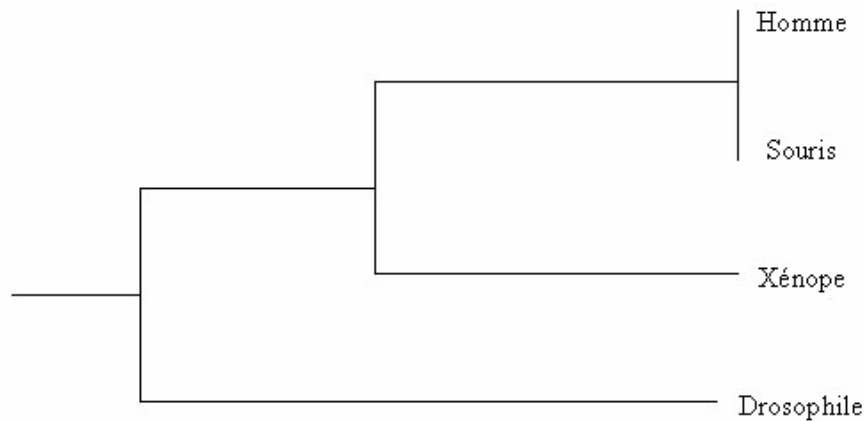
	DP-B7 Homme	DP-B7 Souris	DP-B7 Xénope	DP-Ubx Drosophile
DP-B7 Homme	100 %			
DP-B7 Souris	98,3 %	100 %		
DP-B7 Xénope	96,7 %	96,7 %	100 %	
DP-Ubx Drosophile	90 %	88,3 %	88,3 %	100 %

Résultat des comparaisons des domaines protéiques codés par les gènes AbdB de Drosophile et Hox-B9 de trois Vertébrés : on obtient 86,7 % d'identité globale.

Le tableau ci-dessous précise les identités entre les séquences prises deux à deux :

	DP-B9 Homme	DP-B9 Souris	DP-B9 Xénope	DP-AbdB Drosophile
DP-B9 Homme	100 %			
DP-B9 Souris	100 %	100 %		
DP-B9 Xénope	95 %	95 %	100 %	
DP-AbdB Drosophile	70 %	70 %	70 %	100 %

Si l'on applique la méthode phénétique au traitement de ces données, on peut préciser les parentés entre les êtres vivants concernés : la plus grande parenté étant entre Souris et Homme, puis avec le Xénope, et enfin avec la Drosophile. On peut alors établir la phylogénie suivante :



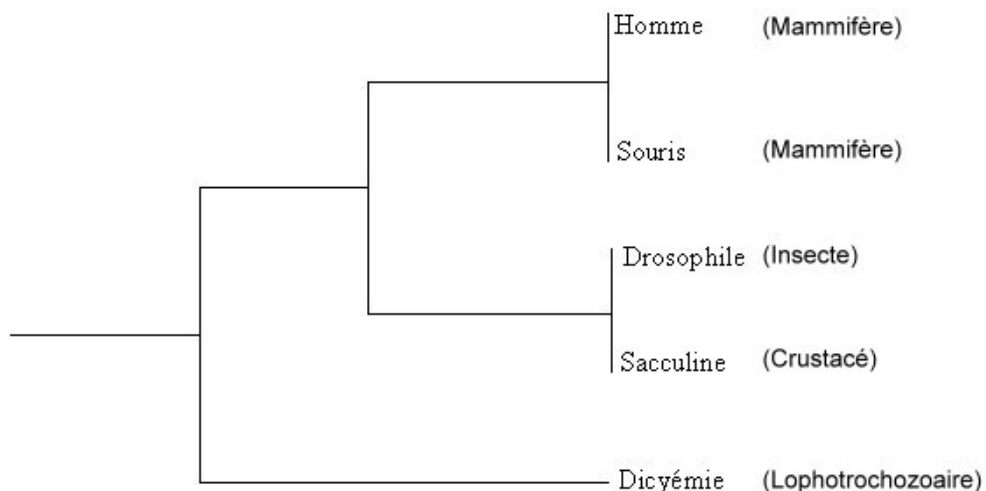
La comparaison des séquences des homéodomaines des gènes Hox B6 de l'Homme et de la Souris avec les gènes Antp de la Drosophile, de la Sacculine et du Dicyemie permet quant à elles d'établir des parentés entre des animaux appartenant à des groupes très éloignés.

Résultat des comparaisons des domaines protéiques codés par les gènes Hox B6 de l'Homme et de la Souris avec ceux codés par les gènes Antp de la Drosophile, de la Sacculine et du Dicyemie : on obtient 71,7 % d'identité globale.

Le tableau ci-dessous précise les identités entre les séquences prises deux à deux :

	DP-B6 Homme	DP-B6 Souris	DP-Antp Drosophile	DP-Antp Sacculine	DP-Antp Dicyemie
DP-B6 Homme	100 %				
DP-B6 Souris	100 %	100 %			
DP-Antp Drosophile	93,3 %	93,3 %	100 %		
DP-Antp Sacculine	93,3 %	73,3,3 %	100 %	100 %	
DP-Antp Dicyemie	73,3 %	73,3 %	75 %	75 %	100 %

L'exploitation des valeurs de ce tableau permet d'établir la phylogénie suivante :



Relations de parenté au sein des Vertébrés – Globines alpha et bêta

Informations scientifiques

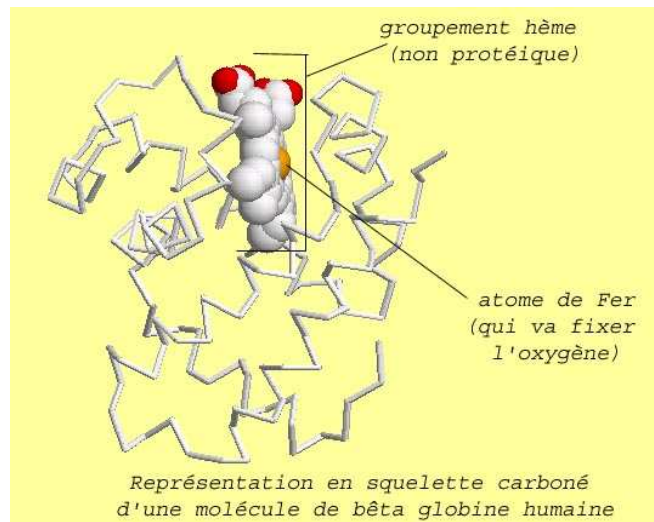
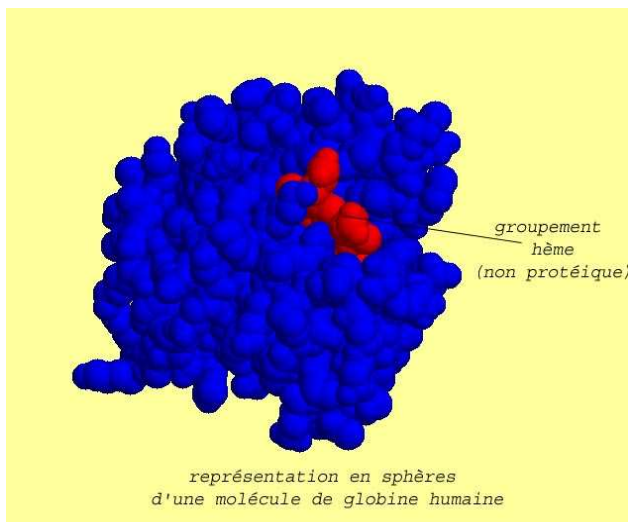
Présentation des globines humaines et des gènes qui les codent

Les globines humaines représentent d'une part les globines qui constituent l'hémoglobine et d'autre part la myoglobine présente dans les cellules musculaires.

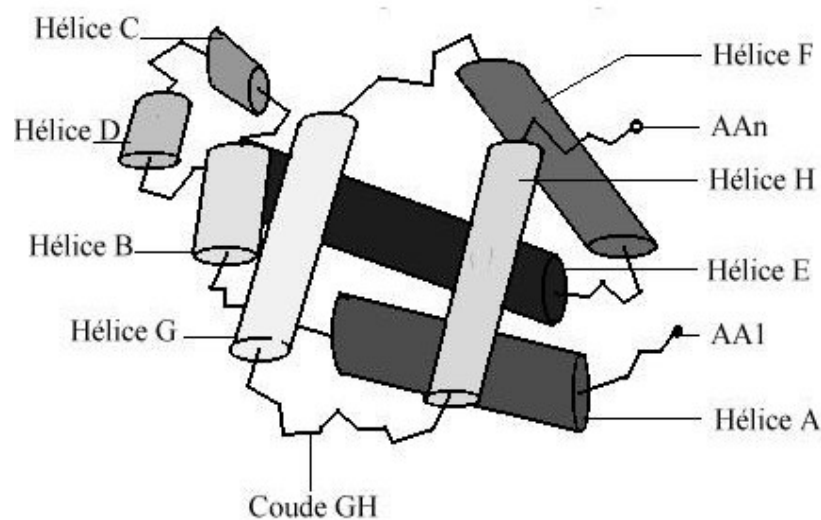
Les globines qui constituent les hémoglobines

Toutes les hémoglobines humaines sont constituées de quatre chaînes polypeptidiques de globines identiques deux à deux.

Chaque molécule d'hémoglobine est un tétramère formé par l'association de quatre chaînes polypeptidiques identiques deux à deux. Chaque chaîne adopte une conformation spatiale lui donnant une forme globuleuse et ménageant une « poche » superficielle dans laquelle se trouve logé le groupement hème (non protéique).

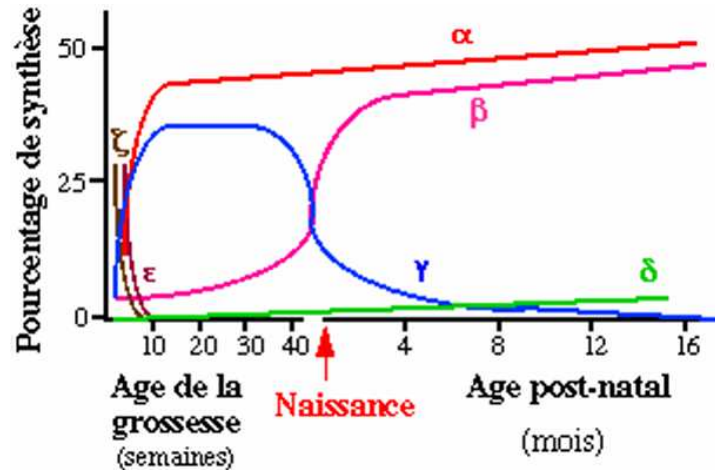


Chaque globine présente sept ou huit secteurs en forme d'hélice droite reliés par des segments comportant parfois des coudes :



Modélisation de la globine bêta humaine

L'être humain, au cours de sa vie, produit plusieurs globines différentes :



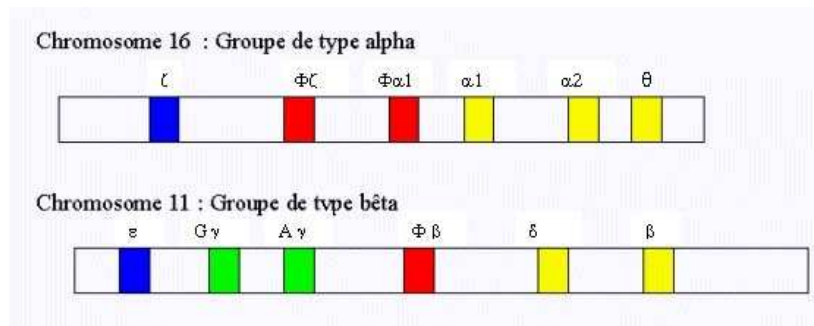
Bien que n'ayant pas la même longueur (141 acides aminés pour les globines alpha 1 et alpha 2, thêta et zêta, 146 acides aminés pour les globines bêta, gamma, delta, epsilon, gamma A et gamma G), toutes ces globines sont fonctionnelles et ont un rôle semblable (fixation et transport du dioxygène grâce au groupement hème).

Avant la naissance, les hématies du fœtus contiennent de l'hémoglobine fœtale constituée de deux chaînes de globine alpha et de deux chaînes de globine gamma. Après la naissance, les hématies contiennent de l'hémoglobine adulte A, très largement majoritaire (97 %), de l'hémoglobine D (2 %) et de l'hémoglobine F (environ 1 %). L'hémoglobine A est constituée de deux chaînes de globine alpha et de deux chaînes de globine bêta. L'hémoglobine D est constituée de deux chaînes de globine alpha et de deux chaînes de globine delta.

Chaque globine est codée par un gène. Les gènes des globines sont répartis sur deux chromosomes différents :

- les gènes de type alpha sont regroupés sur un chromosome (le chromosome numéro 16) ;
- les gènes de type bêta sur un autre (le chromosome numéro 11).

La disposition des gènes de la globine humaine le long des chromosomes correspond à l'ordre dans lequel ils sont exprimés au cours du développement.



Cartes de la famille des gènes codant les chaînes de type alpha et bêta de l'hémoglobine humaine

Les gènes embryonnaires sont représentés en bleu, les gènes foetaux en vert, les gènes adultes en jaune et les pseudogènes en rouge :

- Sur le chromosome 16 on trouve de gauche à droite les gènes **zêta** en bleu, **phi zeta** et **phi alpha 1** en rouge et **alpha 1** puis **alpha 2** et **thêta** en jaune ;
- Sur le chromosome 11, on trouve de gauche à droite les gènes **epsilon** en bleu, **Gamma G** et **gamma A** en vert, **phi bêta** en rouge, **delta** et **bêta** en jaune.

Les pseudogènes ne codent pas pour des protéines fonctionnelles ; la maturation de l'ARNm ou sa traduction sont bloquées.

Tous les Vertébrés à l'exception des Lamproies et Myxines possèdent au moins deux gènes de globine : le gène bêta et le gène alpha. Les protéines d'alpha globine et de bêta globine constituant les molécules d'hémoglobine ont la même fonction chez tous : la fixation et le transport de dioxygène. Toutes les molécules de bêta globine des Vertébrés sont homologues ; il en est de même des molécules d'alpha globine.

Pistes d'exploitation pédagogique des données fournies

En classe terminale, les parentés entre les Vertébrés peuvent être établies à partir des séquences nucléiques et/ou protéiques.

Les données fournies permettent de préciser les parentés entre quelques Vertébrés appartenant à différents groupes à partir de données moléculaires. Ces activités de construction de relations de parenté à partir de données moléculaires viennent en complément des activités de construction d'arbres phylogénétiques à partir de données anatomiques et/ou morphologiques. L'exploitation des données moléculaires permet :

- de confirmer les relations de parenté entre les Vertébrés établies à partir des autres données ;
- de préciser des relations de parenté lorsqu'il y a discussion (par exemple en ce qui concerne les parentés entre Oiseaux et Reptiles).

Séquences et documents

Fichiers des séquences

Dans la banque de thèmes d'étude, le chemin **Parenté entre êtres vivants actuels et fossiles - Phylogénèse, évolution/Relations de parenté au sein des Vertébrés/Globines alpha et bêta permet d'atteindre :**

- **Globine alpha Vertébrés** qui charge le fichier globines-alpha-vertebres.edi affichant les séquences protéiques des alpha globines de douze Vertébrés - Homme, Bonobo, Gorille, Chien, Poule, Canard, Alligator, Caïman, Crocodile, Triton, Thon, Raie ;

- **Globine bêta Vertébrés** qui charge le fichier globines-beta-vertebres.edi affichant les séquences protéiques des bêta globines de douze Vertébrés - Homme, Bonobo, Gorille, Chien, Poule, Canard, Alligator, Caïman, Crocodile, Triton, Thon, Raie.

Documents fournis

Dans la banque de documents, le chemin **Parenté entre êtres vivants actuels et fossiles - Phylogénèse, évolution/Relations de parenté au sein des Vertébrés/Globines alpha et bêta** permet d'atteindre **Globines en 3D** qui charge le fichier presentationhb.jpg présentant les molécules de globine alpha et bêta.

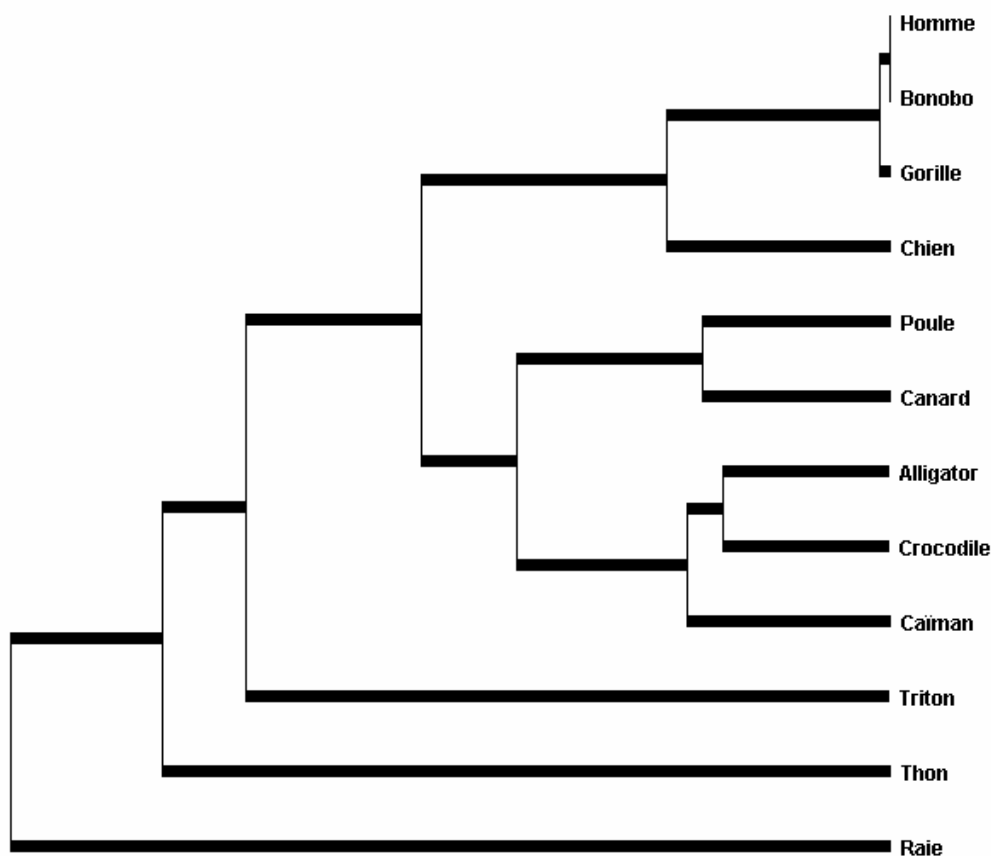
La comparaison des séquences fournies par alignement avec discontinuité permet de remplir une matrice des similitudes. Les résultats et conclusions obtenues à partir des données fournies sont présentés ci-dessous.

À partir de la comparaison des alpha globines de quelques Vertébrés

(fichier globines-alpha-vertebres.edi)

	Homme	Bonobo	Gorille	Chien	Poule	Canard	Alligator	Caïman	Crocodile	Triton	Thon	Raie
Homme	100,0											
Bonobo	100,0	100,0										
Gorille	99,3	99,3	100,0									
Chien	83,7	83,7	84,4	100,0								
Poule	70,2	70,2	69,5	66,7	100,0							
Canard	71,6	71,6	70,9	70,2	87,2	100,0						
Alligator	66,7	66,7	66,0	63,8	71,6	74,5	100,0					
Caïman	67,4	67,4	66,7	62,4	73,0	72,3	85,8	100,0				
Crocodile	68,8	68,8	68,1	66,0	75,9	78,0	87,9	85,1	100,0			
Triton	57,4	57,4	57,4	53,2	56,0	55,3	53,9	54,6	56,0	100,0		
Thon	54,5	54,5	54,5	52,4	49,7	51,0	53,1	51,7	51,7	45,5	100,0	
Raie	45,4	45,4	45,4	45,4	41,1	41,1	44,0	44,0	42,6	37,6	37,6	100,0

Matrice des identités (en pourcentage)



Arbre de parenté que l'on peut obtenir à partir de l'exploitation de cette matrice

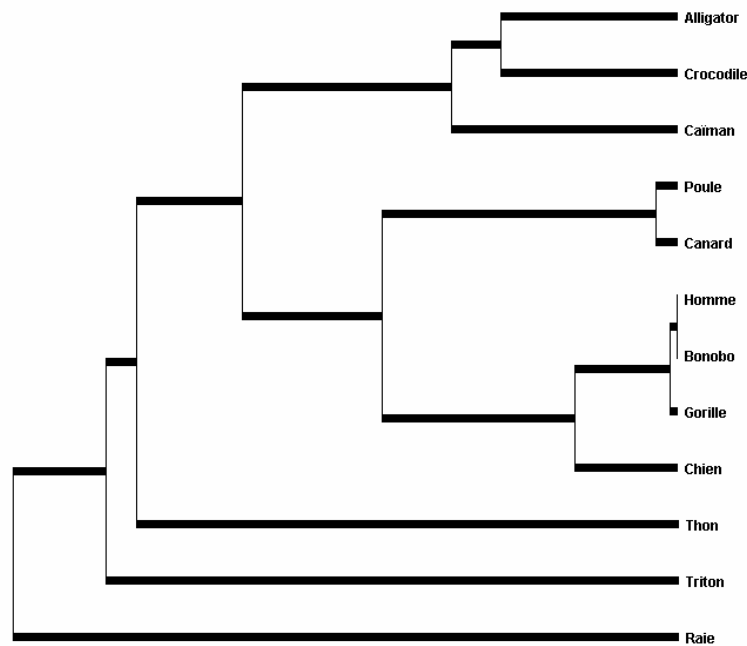
On constate que cet arbre phylogénétique est conforme à celui obtenu à partir de données morphologiques et anatomiques. Aux nœuds successifs de cet arbre, on trouve les ancêtres communs aux Gnathostomes, Ostéichthyens, Tétrapodes, Amniotes, Mammifères et Archosauriens.

À partir de la comparaison des bêta globines de quelques Vertébrés

(fichier globines-beta-vertébres.edi)

	Homme	Bonobo	Gorille	Chien	Poule	Canard	Alligator	Caïman	Crocodile	Triton	Thon	Raie
Homme	100,0											
Bonobo	100,0	100,0										
Gorille	99,3	99,3	100,0									
Chien	89,7	89,7	90,4	100,0								
Poule	69,2	69,2	68,5	71,2	100,0							
Canard	69,2	69,2	68,5	71,2	97,9	100,0						
Alligator	48,6	48,6	49,3	50,7	61,6	62,3	100,0					
Caïman	49,3	49,3	50,0	52,7	57,5	58,2	79,5	100,0				
Crocodile	54,8	54,8	55,5	55,5	63,7	64,4	80,8	72,6	100,0			
Triton	46,9	46,9	47,6	46,9	46,9	47,6	44,1	42,8	42,8	100,0		
Thon	47,9	47,9	47,3	50,0	54,8	54,8	41,8	41,1	41,1	41,1	100,0	
Raie	38,7	38,7	38,7	40,1	42,3	40,8	33,8	33,1	34,5	30,3	38,0	100,0

Matrice des identités (en pourcentage)



Arbre de parenté que l'on peut obtenir à partir de l'exploitation de cette matrice

L'arbre obtenu à partir des chaînes de globine bêta des mêmes Vertébrés montre certaines des parentés révélées par l'arbre précédent (tous les Mammifères groupés ensemble, Canard et Poule dans le même groupe par exemple). Mais il indique aussi deux différences importantes :

- le Thon est plus étroitement apparenté aux Amniotes que le Triton ;
- les Oiseaux sont plus étroitement apparentés aux Mammifères qu'aux Crocodiliens.

Non seulement l'arbre obtenu à partir des globines bêta est différent de celui des globines alpha, mais de plus, il n'est pas conforme aux données morphologiques et anatomiques. Certaines des informations qu'il traduit sont considérées comme erronées. Cela permet de souligner la limite des informations phylogénétiques extraites des données moléculaires surtout avec la technique d'analyse utilisée au lycée qui est phénétique, c'est-à-dire basée sur le nombre de différences ou de ressemblances entre les molécules. Pour que les informations soient totalement fiables, il faudrait que la vitesse d'évolution des molécules homologues (mutations et fixation des mutations) dans les différentes lignées soit exactement la même. Ce n'est jamais le cas. Lorsque les divergences des lignées sont éloignées dans le temps, il est possible qu'une vitesse plus grande dans une lignée conduise à des résultats inattendus.

COMPLÉMENTS

D'autres fichiers comportant un plus grand nombre de séquences sont fournis. Ils peuvent être utilisés pour créer des sous-ensembles et pour discuter des informations qu'ils fournissent par comparaison à une phylogénie reconnue construite à partir de données morphologiques et anatomiques. Cela peut être utile aussi pour faire travailler des groupes d'élèves sur des exemples différents.

Fichiers des séquences

Dans la banque de thèmes d'étude, le chemin **Parenté entre êtres vivants actuels et fossiles - Phylogénèse, évolution/Relations de parenté au sein des Vertébrés/Globines alpha et bêta** permet d'atteindre :

- **Globines alpha 10 Vertébrés** qui charge le fichier globinesalphareptiles.edi affichant les séquences protéiques des alpha globines de dix Vertébrés, dont sept « Reptiles » (Cobra, Vipère, Iguane, Tortue, Crocodile, Alligator, Caïman), deux Oiseaux (Poule, Canard) et un Batracien (Triton) ;
- **Divers alpha globines** qui charge le fichier alphaglobinetotal.edi affichant les séquences protéiques de l'alpha globine chez de nombreux Vertébrés appartenant à des classes différentes : Mammifères (Homme, Bonobo, Gorille, Panda, Dauphin, Rorqual, Chien, Otarie, Bœuf, Yak, Lama, Souris), Oiseaux (Manchot, Poule, Canard), Reptiles Chéloniens (Tortue géante), Lépidosauriens (Cobra, Vipère, Liophis, Iguane), Crocodiliens (Crocodile, Alligator, Caïman), Poissons Ostéichthyens (Poisson zèbre, Thon), Poissons Chondrichthyens (Raie) ;
- **Divers bêta globines** qui charge le fichier betaglobinetotal.edi affichant les séquences protéiques de la bêta globine chez de nombreux Vertébrés appartenant à des classes différentes : Mammifères (Homme, Chimpanzé, Gorille, Lemur, Rorqual, Dauphin, Chien, Panda), Oiseaux (Poule, Canard, Cormoran), Reptiles Lépidosauriens (Liophis, Varan), Crocodiliens (Crocodile, Alligator, Caïman), Chélonien (Tortue carette), Amphibiens (Triton, Crapaud, Xénope), Poissons Ostéichthyens (Poisson rouge, Thon), Poissons Chondrichthyens (Raie).

Relations de parenté au sein des Primates – Gène de l'opsine S

Informations scientifiques

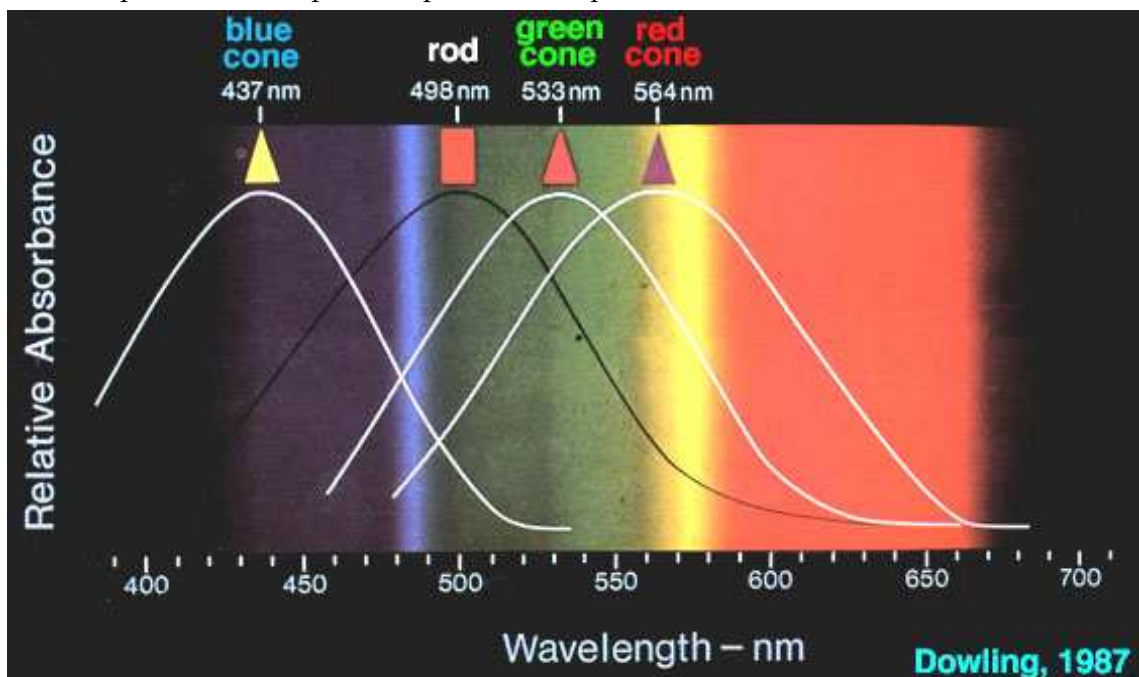
La rétine humaine est un tissu nerveux qui renferme notamment des cellules photoréceptrices de deux types : les cônes et les bâtonnets. Les cônes sont responsables de la vision des couleurs ; les bâtonnets ne permettent pas la vision des couleurs, mais sont très sensibles à la lumière et permettent une vision en très faible éclaircissement.

La vision des couleurs est liée à l'existence dans la rétine de trois types cellulaires (les cônes) qui sont des photorécepteurs synthétisant trois pigments différents, des opsines, qui présentent des niveaux d'absorption différents dans le rouge (opsine rouge ou L), le vert (opsine verte ou M) ou le bleu (opsine bleue ou S). Avec un seul des pigments, on ne voit pas les couleurs ; la présence des trois pigments permet une vision trichromatique.

Les pigments des cellules photoréceptrices

Toutes ces cellules contiennent un pigment, de nature essentiellement protéique, excitable par des photons. Ce pigment varie selon le type de cellules photoréceptrices :

- les bâtonnets synthétisent un pigment appelé **rhodopsine** (constituée d'une opsine et d'un groupement prosthétique, le 11-cis-rétinal lié à la lysine 216 de la protéine, dont le précurseur est le all-trans-rétinol ou vitamine A. Un déficit alimentaire en vitamine A conduit ainsi à l'héméralopie ou cécité nocturne). La rhodopsine présente une bande d'absorption large dans la région visible du spectre avec un pic à 500 nm.
- les cônes synthétisent un pigment appelé opsine ; il existe trois types de cônes en fonction du pigment qu'ils synthétisent : les cônes à opsine S (Short ou bleue), M (Médium ou verte) et L (Long ou rouge), ces différentes opsines différant par leur spectre d'absorption.



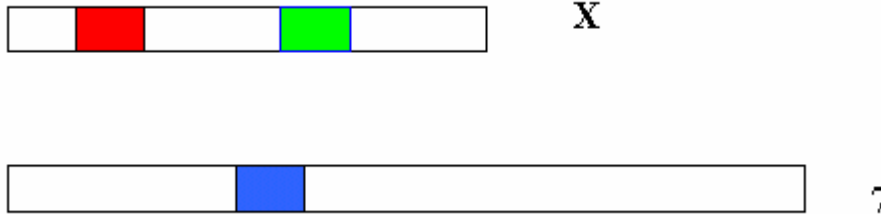
Les spectres d'absorption des trois opsines photoréceptrices ont été mesurés en éclairant des cônes avec un faisceau de lumière de 1 mm seulement. Les réponses des différents cônes montrent trois groupes : certaines sont excitées au maximum par la lumière bleue, d'autres sont excitées au maximum par la lumière verte et les dernières par la lumière orangée-rouge.

Les cônes S contiennent une opsine ayant un maximum d'absorption dans les courtes longueurs d'onde du visible (« opsine S ou bleue »), les cônes M contiennent une opsine ayant un maximum d'absorption dans les longueurs d'onde moyennes du visible (« opsine M ou verte »), les cônes L contiennent une opsine ayant un maximum d'absorption dans les grandes longueurs d'onde du visible (« opsine L ou rouge »). Bien entendu, tous les photorécepteurs possèdent les mêmes gènes codant pour les opsines, mais un seul de ces gènes s'exprime dans un photorécepteur donné.

Les protéines photoréceptrices des cellules en cône sont, comme l'opsine de la rhodopsine, des récepteurs à sept hélices, et elles contiennent elles aussi le 11-cis-rétinal.

Localisation chromosomique des gènes codant pour les opsines humaines :

Les gènes codant pour les opsines M ou verte et L ou rouge sont situés sur le chromosome X. Le gène codant pour l'opsine S ou bleue est porté par le chromosome 7.



Les opsines chez les Primates

Chez les Primates, la vision trichromatique n'est possible que dans certains groupes :

- les Singes de l'Ancien Monde (Afrique, Asie et Europe), ou catarrhiniens, possèdent, comme l'Homme :
 - le gène de l'opsine S (bleue) porté par un autosome, codant pour une opsine présentant un pic d'absorption à 435 nm ;
 - le gène de l'opsine M (verte) porté par le chromosome X, codant pour une opsine présentant un pic d'absorption à 535 nm ;
 - le gène de l'opsine L (rouge) porté par le chromosome X, codant pour une opsine présentant un pic d'absorption à 565 nm.

Ces Singes sont donc tous trichromates ;

- les Singes du Nouveau Monde (Amérique), ou platyrrhiniens, possèdent :
 - le gène de l'opsine S (bleue), situé sur un autosome et codant pour une opsine ayant un pic d'absorption à 430 nm environ ;
 - un seul autre gène codant pour une opsine S/M, porté par le chromosome X, codant pour une opsine ayant un pic d'absorption de 530 à 560 nm. Or il existe trois allèles différents possibles pour cette opsine, codant pour des pigments dont les pics d'absorption se situent vers 536 nm, 556 nm, 563 nm.

Chez ces Singes, comme le Saïmiri par exemple (Singe écureuil), on a ainsi plusieurs phénotypes possibles pour la vision des couleurs (Singes dichromates ou Singes trichromates), notamment en fonction du sexe :

- les mâles sont tous dichromates : ils possèdent le gène de l'opsine S, et un seul gène pour une opsine M/L car ils n'ont qu'un seul chromosome X ;
- les femelles, qui possèdent deux chromosomes X, peuvent être dichromates ou trichromates : elles ont toutes un gène de l'opsine S, et possèdent un seul allèle du gène de l'opsine M/L si elles sont homozygotes, mais en possèdent deux si elles sont hétérozygotes. Si les deux allèles qu'elles possèdent sont différents, elles produiront donc trois opsines différentes, ce qui permettra une vision trichromatique.

On a d'autre part récemment découvert que les Singes hurleurs (*Alouata seniculus*) étaient trichromates, aussi bien les mâles que les femelles. Ils synthétisent trois pigments visuels présentant des pics d'absorption à 430 nm, 532 nm et 562 nm. C'est actuellement le seul Singe du Nouveau Monde pour lequel on ait découvert cette trichromatie.

- les Lémuriens : ceux qui vivent la nuit ont une vision en noir et blanc ; ceux qui vivent le jour ont tous une vision dichromatique. Ils possèdent le gène de l'opsine S (bleue) sur un autosome, et le gène d'une opsine M/L sur le chromosome X.

Pistes d'exploitation pédagogique des données fournies

Il est possible d'établir des relations de parenté entre les Primates à l'aide des données moléculaires concernant les opsines, et notamment l'opsine S (ou opsine bleue). Tous les Primates possèdent ce gène, porté par un autosome, et il code pour une opsine ayant toujours un pic d'absorption voisin de 430 nm.

Séquences et documents

Fichiers des séquences

Dans la banque de thèmes d'étude, le chemin **Parenté entre êtres vivants actuels et fossiles - Phylogénèse, évolution/Relations de parenté au sein des Primates/Gène de l'opsine S (Shortwave - Bleu)** permet d'atteindre **Opsine S bleu Primates** qui charge le fichier OpsPrimates.edi affichant les séquences protéiques des opsines bleues (ou opsines S) de quelques Primates appartenant à des groupes différents :

- platyrrhiniens : Cebus (Capucin), Saïmiri (Singe écureuil), Alouate (Singe hurleur) ;
- catarrhiniens : Homme, Macaque, Bonobo, Chimpanzé, Gorille.

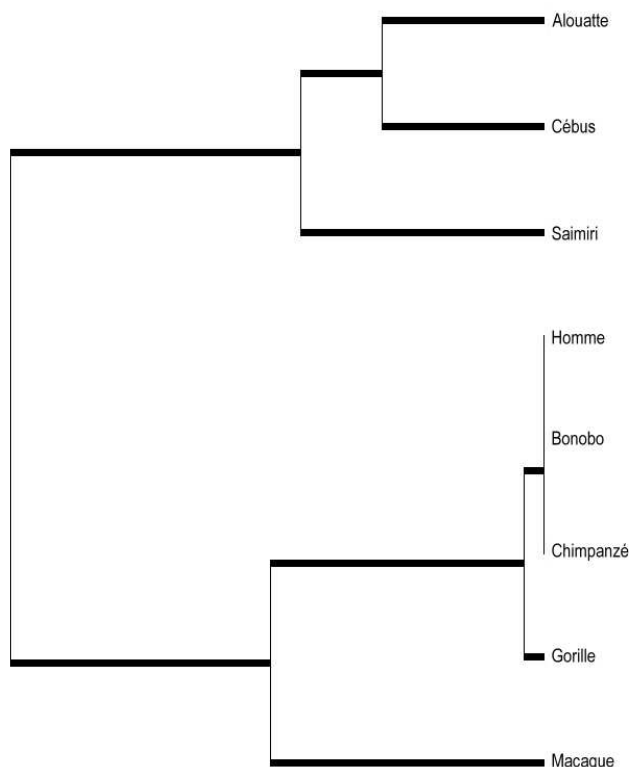
Documents fournis

Dans la banque de documents, le chemin **Parenté entre êtres vivants actuels et fossiles - Phylogénèse, évolution/Relations de parenté au sein des Primates/Gène de l'opsine S (Shortwave - Bleu)** permet d'atteindre **Opsine S** pour afficher le fichier Opsine S.bmp présentant la nature et le rôle de l'opsine S ainsi que la localisation du gène qui la code.

Les informations tirées de la comparaison des séquences par alignement avec discontinuités permettent de construire la matrice des identités suivante (les valeurs sont en pourcentage) :

	Homme	Bonobo	Chimpanzé	Gorille	Macaque	Alouate	Saïmiri	Cebus
Homme	100							
Bonobo	100	100						
Chimpanzé	100	100	100					
Gorille	99,7	99,7	99,7	100				
Macaque	96	96	96	95,7	100			
Alouate	92,3	92,3	92,3	92	92,3	100		
Saïmiri	91,7	91,7	91,7	91,4	92,2	97,1	100	
Cebus	92,5	92,5	92,5	92,2	92,5	97,7	96	100

À partir de cette matrice, on peut établir la phylogénie suivante :



Relations de parenté au sein des Hominidés - ADN mitochondrial

D'autres informations sur ce thème peuvent être trouvées sur le site *Vie à Jussieu* dont l'adresse est : <http://www.snv.jussieu.fr/vie/documents/adnancien/adnmt.htm>

Informations scientifiques

ADN fossile et histoire des Homininés

Un arbre phylogénétique des Homininés basé sur l'utilisation de caractères morphologiques et anatomiques peut être construit avec le logiciel *Phylogène*.

Les données moléculaires commencent à être utilisées pour aborder certains problèmes de l'histoire évolutive des Homininés. Sous ce terme, on désigne ici les espèces appartenant à la lignée humaine qui partagent avec *Homo Sapiens* un ancêtre commun qui n'est pas celui des Chimpanzés. L'espèce humaine est la seule espèce d'Homininés qui existe aujourd'hui. Pour exploiter les données moléculaires relatives à d'autres espèces, il faut se procurer de l'ADN de ces espèces, donc recourir à de l'ADN fossile. La première étude date de 1997 et elle est relative à l'Homme de Neandertal.

Après la mort d'un organisme, l'ADN se dégrade rapidement, surtout si la température de la région est assez élevée. L'ADN que l'on peut extraire d'un fossile se trouve sous forme de courts fragments de quelques dizaines à une ou deux centaines de nucléotides. Au-delà de 100 000 ans, il est illusoire de vouloir récupérer de l'ADN utilisable pour déchiffrer une séquence. Cela réduit l'utilisation de l'ADN fossile aux périodes les plus récentes de l'histoire de la lignée humaine, donc aux « espèces » qui sont contemporaines des premiers Sapiens. Pour l'instant, on n'a utilisé que de l'ADN de neandertaliens. *A priori*, il pourrait être possible d'exploiter l'ADN de l'Homme de Flores qui a vécu jusqu'à -12 000 ans, bien que les conditions climatiques de l'Indonésie ne soient pas favorables à la préservation de l'ADN. Si cela s'avérait réalisable, les informations obtenues pourraient éclairer la place controversée de cet Homininé dans l'histoire de la lignée humaine.

Intérêt de l'ADN mitochondrial

Les études sur l'ADN fossile des Homininés ont porté sur l'ADN mitochondrial et non sur l'ADN nucléaire. Cela s'explique pour diverses raisons :

- presque tout l'ADN disparaît au cours de la fossilisation. Il n'en reste que des courts fragments. Le premier travail consiste à amplifier par la technique de la PCR (polymerase chain reaction) un ou plusieurs fragments qui ont été préservés. Il faut choisir d'avance ceux que l'on recherche dans l'extrait fossile. Il y a généralement plusieurs centaines de mitochondries par cellule et chacune contient une molécule circulaire d'ADN de plus de 16 000 paires de bases. Les séquences d'ADN mitochondrial sont donc au nombre de plusieurs centaines dans une cellule alors que celles de l'ADN nucléaire sont au nombre de deux sauf pour les séquences répétées. Il est donc **plus probable de trouver dans l'ADN fossile des fragments d'ADN mitochondrial** susceptibles d'être amplifiés que des fragments d'ADN nucléaire ;
- un autre point important est que l'ADN mitochondrial évolue plus rapidement que l'ADN nucléaire : les mutations s'y fixent plus rapidement. Pour des périodes relativement courtes imposées par la préservation de l'ADN fossile, l'ADN mitochondrial est donc **plus informatif** que l'ADN nucléaire. En outre, le génome mitochondrial ne se transmet que de la mère à l'enfant (les mitochondries du spermatozoïde ne contribuent pas au stock initial de mitochondries du zygote). Il n'y a donc pas de recombinaison. Les différences entre molécules d'ADN mitochondrial résultent uniquement de mutations ;
- le séquençage de l'ADN mitochondrial d'un très grand nombre d'Hommes actuels a révélé que deux régions appartenant à la zone de contrôle de la réplication de cet ADN présentent une grande variabilité : ce sont les régions hypervariables 1 et 2. Ce sont des fragments de ces régions hypervariables qui ont été recherchés chez les neandertaliens.

Objectifs de l'étude de l'ADN mitochondrial fossile

Les recherches sur l'ADN mitochondrial des neandertaliens ont pour principal objectif de déterminer si ces Homininés ont pu contribuer au pool génétique des populations de *Homo sapiens*.

Cette question s'insère dans le problème plus général de l'origine des Hommes modernes. Au cours des années 1980, deux hypothèses extrêmes se sont opposées. Selon l'hypothèse multirégionale, dite encore du candélabre, des populations d'*Homo erectus* d'origine africaine se sont répandues à travers le monde et ont graduellement et indépendamment évolué pour donner d'abord des *Homo Sapiens* archaïques, puis des Hommes modernes. Selon cette hypothèse les Hommes de Neandertal seraient issus d'*Homo Sapiens* archaïques européens (les préneandertaliens) et

seraient à l'origine des *Homo sapiens* européens et du Proche-Orient. Selon la seconde hypothèse, dite de l'arche de Noé, tous les *Homo sapiens* actuels auraient pour origine une population initiale du proche orient ou d'Éthiopie qui aurait migré ensuite dans tout le monde. Cette population résulterait d'un processus de spéciation, il y a 150 000 ans environ. Elle se serait ensuite répandue dans toutes les régions du monde sans se métisser avec les populations d'*Homo sapiens* archaïques déjà présentes. Ces *Homo sapiens sapiens* auraient contribué par des mécanismes mal élucidés à la disparition des *Homo sapiens* archaïques, des neandertaliens en Europe.

Aujourd'hui, il y a un accord assez général pour dire que la première hypothèse, celle du candélabre ne s'applique pas à l'Europe. En effet, on ne connaît pas de fossiles d'*Homo* permettant d'asseoir l'hypothèse d'une transformation graduelle de l'Homme de Neandertal en *Homo sapiens*. Toutefois, au Proche-Orient, les populations de neandertaliens ont cohabité pendant plus de 50 000 ans avec celles des premier *Homo sapiens* ; en Europe, la cohabitation a duré plus de 10 000 ans (de -40 000 ans à -29 000 ans environ, date de disparition des derniers neandertaliens). Au cours de cette période de cohabitation, on ne peut exclure des possibilités de métissage entre les Hommes de Cro-Magnon et les Hommes de Neandertal. Si tel est le cas, l'Homme de Neandertal appartiendrait à la même espèce que l'Homme actuel et devrait s'appeler *Homo sapiens neandertalensis*. En revanche, s'il n'y a pas eu hybridation, l'Homme de Neandertal serait une espèce différente de celle de *Homo sapiens* et mériterait le nom de *Homo neandertalensis*. Dans ce cas, les neandertaliens représenteraient un cul-de-sac évolutif puisque disparus sans laisser de descendance.

C'est à la résolution de cette question, métissage ou non entre neandertaliens et premiers Hommes modernes, que visent les études sur l'ADN mitochondrial.

Apports des séquences d'ADN mitochondrial

C'est en 1997 qu'a été déterminée la première séquence d'ADN mitochondrial d'un neandertalien. Il s'agit de la séquence, longue de 379 paires de nucléotides, de la région hypervariable 1 de la zone où se trouve l'origine de la réplication de l'ADN. Elle a été établie à partir d'un extrait d'ADN du spécimen type de l'Homme de Neandertal, celui de Feldhofer (vallée de Neander, Allemagne) découvert en 1856. Elle a été comparée à 994 séquences homologues d'Hommes actuels appartenant à différentes régions du monde dont celles de 510 Européens. Le nombre moyen de différences entre les séquences des Hommes actuels est de huit. Le nombre moyen de différences entre la séquence de l'Homme de Feldhofer et celles des Hommes modernes est de 27, soit plus de trois fois plus importante. La plus petite différence entre la séquence neandertalienne et celle d'un Homme actuel est de 22 et la plus grande de 36. La séquence neandertalienne se place donc nettement au-delà de la variabilité constatée entre Hommes actuels. Le nombre de différences entre la séquence neandertalienne et celles des Asiatiques et des Africains est similaire. En particulier, aucune séquence d'Européen ne s'est avérée plus proche de celle de l'Homme de Feldhofer. Ces données n'apportent aucun argument en faveur de l'hypothèse suivant laquelle les neandertaliens auraient contribué au pool génétique des Européens par suite de métissage avec les Hommes de Cro-Magnon, en réalité le métissage entre femmes neandertaliennes et Hommes de Cro-Magnon, puisque l'ADN mitochondrial se transmet uniquement par les femmes.

Trois autres séquences de la région hypervariable 1 ont par la suite été déterminées à partir de fossiles neandertaliens d'âges différents et appartenant à des régions de l'Europe très éloignées les unes des autres : Allemagne, Caucase, Croatie. Elles ont entièrement confirmé les indications fournies par la première séquence. De plus, quand on établit un arbre phylogénétique à partir de ces séquences et de celles des Hommes actuels, on constate que les séquences des neandertaliens se regroupent entre elles ; les neandertaliens paraissent avoir un ancêtre commun qui n'est pas celui des Hommes modernes. Le nombre moyen de différences entre ces séquences neandertaliennes appartenant à des individus ayant vécu dans des régions européennes éloignées et à des moments différents (-40 000 à -29 000 ans) est du même ordre que celles constatées entre les séquences des Hommes actuels.

Si tout cela ne plaide pas en faveur du métissage entre neandertaliens et Hommes de Cro-Magnon, on ne peut pas écarter l'hypothèse qu'il a eu lieu mais que l'on n'en retrouve pas de signe dans les populations européennes actuelles. En effet, les descendants de ces mélanges ont pu appartenir à des populations qui n'ont pas survécu au cours du paroxysme de la dernière période glaciaire. On pouvait donc penser que l'on avait plus de chance de repérer ce métissage dans l'ADN fossile des Hommes de Cro-Magnon qui ont cohabité avec les neandertaliens.

En mai 2003, l'équipe de Caramelli a publié la séquence de la région hypervariable 1 de deux Hommes de Cro-Magnon ayant vécu en Italie, il y a 24 000 ans. Leurs séquences sont comparables à celles des Hommes actuels ; dans un arbre phylogénétique, ces Hommes de Cro-Magnon se retrouvent parmi les Hommes actuels et non avec les neandertaliens.

En 2004, l'équipe de David Serre et d'André Langaney a publié les résultats d'une étude sur l'ADN fossile de neandertaliens (La Chapelle-aux-Saints, Engis) et d'Hommes de Cro-Magnon (cinq spécimens, notamment des sites

de La Madeleine et de Cro-Magnon). Disposant des informations sur les séquences de neandertaliens précédentes, ils ont essayé d'obtenir sur les différents échantillons d'ADN fossile l'amplification de fragments de la région hypervariable 1 par la technique PCR en utilisant des amorces typiques des séquences neandertaliennes. Ils ont obtenu des produits d'amplification pour tous les extraits de neandertaliens mais aucun à partir des extraits de Cro-Magnon.

Ainsi, jusqu'ici, aucune trace neandertalienne n'a été mise en évidence chez les Hommes de Cro-Magnon. Il semble que l'Homme de Neandertal n'a légué aucun tribut génétique à l'Homme moderne.

Pistes d'exploitation pédagogique des données fournies

Séquences et documents

Fichiers des séquences

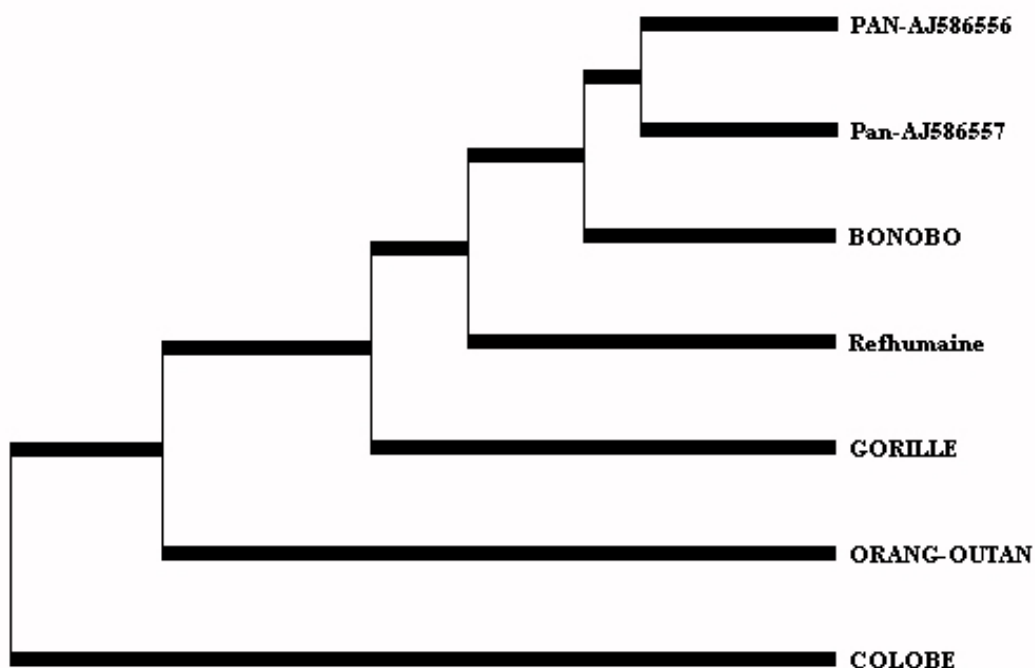
Dans la banque de thèmes d'étude, le chemin **Parenté entre êtres vivants actuels et fossiles - Phylogénèse, évolution/Relations de parenté au sein des Hominidés/ADN mitochondrial** permet d'atteindre :

- **Primates** qui charge le fichier Primates.edi affichant la séquence consensus humaine de référence et de divers Primates ;
- **Actuels et fossiles** qui charge le fichier 7mod+3neand+2cromagnon-2pan.edi affichant les séquences de sept Hommes actuels appartenant à divers continents, trois neandertaliens (deux de Feldhofer, l'autre de Croatie), deux Hommes de Cro-Magnon (ceux d'Italie) et deux Chimpanzés ;
- **ADN mitochondrial partiel** qui charge le fichier ADN-Mito-partiel.edi portant sur une portion de cette séquence de 345 nucléotides, et qui a été élaboré à partir des séquences d'un fragment d'amplification de neandertaliens obtenus par l'équipe de David Serre et d'André Langaney, les affichant pour l'Homme moderne, Cro-Magnon, neandertalien.

Les séquences fournies sont celles de la région hypervariable 1 de l'ADN mitochondrial de divers Primates et d'Homininés. Afin de pouvoir faire des comparaisons, on a sélectionné des séquences homologues de 345 nucléotides à partir des séquences fournies par la banque internationale sur la région hypervariable 1.

Bien entendu, l'exploitation de ces fichiers nécessite un minimum d'informations préalables sur l'ADN mitochondrial et l'ADN fossile. L'exploitation des données fournies est de même nature que celle des séquences utilisées précédemment pour établir des phylogénies.

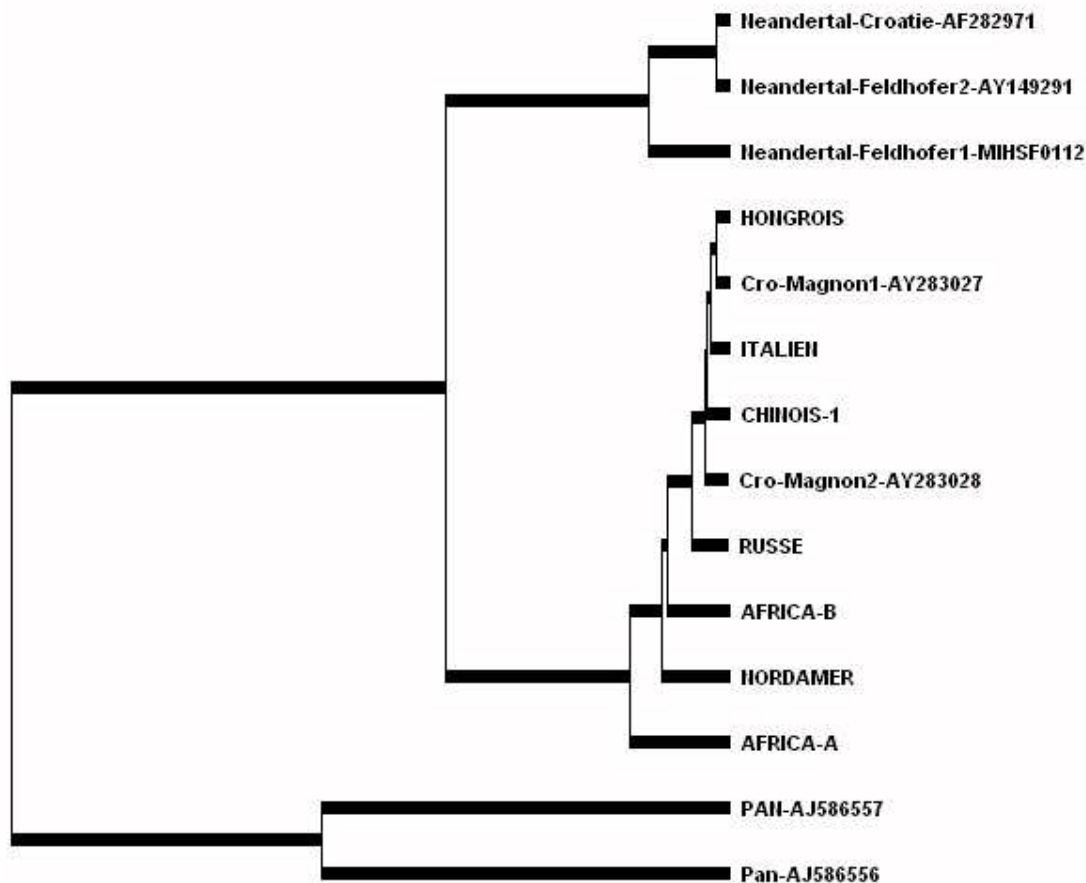
Le fichier sur les Primates permet de discuter la pertinence d'utiliser l'ADN mitochondrial pour établir des relations de parenté. Le principe est de comparer les séquences deux à deux à l'aide de la fonction de comparaison d'*Anagène* afin de construire la matrice des ressemblances entre les régions hypervariables 1 des divers Primates. À partir de là, on peut construire un arbre phylogénétique des Primates similaire à celui obtenu avec *Phylogène*.



NB : Pan correspond au Chimpanzé.

Cet arbre est conforme à celui construit en utilisant d'autres molécules : l'ADN mitochondrial peut donc être utilisé pour reconstruire des phylogénies. De plus, on constate que par rapport aux séquences d'ADN nucléaire utilisées, le nombre de différences entre les séquences d'ADN mitochondrial de l'Homme et des autres Hominidés (Chimpanzé, Bonobo, Gorille) est relativement important, ce qui indique une évolution rapide de cette région de l'ADN mitochondrial. Elle peut donc être utilisée pour résoudre des problèmes de parenté au sein des Homininés.

Le second fichier (Homme moderne, Cro-Magnon, neandertalien, Chimpanzé) peut être utilisé pour traiter la question de la place de l'Homme de Neandertal par rapport à *Homo sapiens* et donc discuter la possibilité ou non d'un métissage entre ces deux *Homo*. Là aussi, il s'agit de comparer les séquences deux à deux afin de construire une matrice des ressemblances et finalement un arbre phylogénétique tel que celui représenté ci-après.



On y voit que les neandertaliens partagent avec *Homo sapiens* un ancêtre commun qui n'est pas celui du Chimpanzé, que les neandertaliens forment un groupe monophylétique à part de celui des Hommes actuels, que les Hommes de Cro-Magnon ne sont pas situés au sein des neandertaliens mais parmi les *Homo sapiens* actuels. Cet ensemble d'informations forme un faisceau d'arguments en faveur de l'absence d'hybridation entre ces deux types d'*Homo* et donc de leur appartenance à deux espèces.

Le troisième fichier des « séquences partielles » présente l'intérêt de fournir des données sur toutes les séquences de neandertaliens ayant été déterminées. Il permet de rechercher si, dans cette courte portion de la région hypervariable 1, les séquences de neandertaliens présentent à certains sites des caractères communs différents de ceux que l'on trouve chez les *Homo sapiens* passés et actuels.

Si on ne veut pas utiliser le deuxième fichier, on peut très bien fournir l'arbre phylogénétique et demander de rechercher en quoi les données du troisième fichier sont en accord avec cet arbre. Il suffit de faire une comparaison avec discontinuité de toutes les séquences partielles.