



Alignements de séquences



Travaux des Actions Académiques Mutualisées

Niveau

- Terminale STL Biotechnologies

Thème du programme

- Cours de CBSV - Chapitre 5.1. Les organismes vivants sont divers mais apparentés.
- Activités Technologiques de Biotechnologies

Situations pédagogiques

- Travaux dirigés

Liens internet

- <http://www.ncbi.nlm.nih.gov>
- <http://blast.ncbi.nlm.nih.gov>
- <http://www.ebi.ac.uk/Tools/msa/clustalw2/>

Compétences B2i

- **Domaine1: s'approprier un espace numérique de travail.**
- **Domaine 3: créer, produire, traiter, exploiter des données.**

Matériels TICE

- Un poste PC par binôme
- Une connexion internet
- Logiciel de traitement de texte et d'images

Mots clés

- Alignement de séquences, pourcentage d'identité, arbre phylogénétique, BLAST, CLUSTALW.



Votre avis nous intéresse, merci de répondre à notre enquête concernant ce scénario.

Elève, cliquer [ici](#).

Professeur, cliquer [ici](#).



Préambule

| | | |
|-----------------------------------|--|---|
| Objectifs | <ul style="list-style-type: none"> Mobiliser les concepts fondamentaux de biologie moléculaire pour compléter un document | |
| Durée | <ul style="list-style-type: none"> 30 min | |
| Ressources complémentaires | <ul style="list-style-type: none"> Annexe 1 : Détermination du taux d'homologie- pourcentage d'identité et score d'alignement |  Cliquer pour télécharger |
| | <ul style="list-style-type: none"> Annexe 2 : Présentation rapide de deux programmes d'alignement de séquence |  Cliquer pour télécharger |
| Liens utilisés | http://www.genie-bio.ac-versailles.fr/IMG/pdf/traam-bioinfo1-annexe1-taux-homologie.pdf http://www.genie-bio.ac-versailles.fr/IMG/pdf/traam-bioinfo1-annexe2-programmes-en-ligne.pdf | |



Questions

- Compléter les phrases du document 1 par les termes scientifiques adaptés.
- Exploiter les informations apportées par l'annexe 1 en donnant, pour la séquence proposée dans l'annexe, le pourcentage d'identité et le score d'alignement.
- A partir de l'annexe 2, expliquer simplement à quoi servent respectivement NCBI, Blast et ClustalW.

Document 1

Séquence d'ADN :

C'est l'ordre d'enchaînement des _____ d'un brin d'ADN. Une séquence est orientée de l'extrémité _____ à l'extrémité _____. Les deux brins d'une molécule d'ADN étant _____, la connaissance de la séquence d'un seul des deux brins est suffisante pour caractériser la molécule en question.

Stabilité et instabilité des séquences d'ADN au cours des générations :

La division cellulaire nécessite la _____ de l'ADN de la cellule mère en deux molécules d'ADN retrouvées dans les cellules filles.

La _____ est un processus extrêmement fidèle qui peut cependant faire des erreurs. Ces erreurs portent le nom de _____ et s'accumulent au fil des divisions (= des générations).

Homologie de séquences:

Homologie = parenté.

On parle d'homologie pour deux séquences ayant évolué à partir du même ancêtre commun. Plus ces séquences se ressemblent et plus elles ont un ancêtre commun _____ dans le temps.

Alignement de séquences :

Un alignement de séquence consiste à comparer deux séquences mises l'une à côté de l'autre.

La comparaison des séquences permet d'observer les régions conservées et d'évaluer le lien de _____ entre les séquences étudiées.

**Activité n°1 : "Inspecteur BLAST et la Bête du Gévaudan"**

| | |
|---------------------|--|
| Durée | 1h |
| Modalités | <i>Rendre le travail sous forme d'un document word (ou autre traitement de texte) avec saisies écrans justificatives. Penser à mettre son nom dans le nom du fichier afin de le reconnaître rapidement.</i> |
| Lien utilisé | BLAST : http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome |

Contexte :

En 1766 près de Villefranche de Rouergue, une jeune fille a été dévorée par la mystérieuse « bête du Gévaudan ». Quelle est la vraie nature de cette « bête » ? Loup ? Chien fou ? Psychopathe tueur en série ? Loup-Garou ?...

Les restes de la jeune victime sont encore conservés comme reliques dans la crypte de l'église. On a pu retrouver du sang séché sous ses ongles. Le laboratoire de la police scientifique a été chargé d'analyser ce sang séché. L'analyse a permis d'identifier et d'amplifier par PCR la séquence suivante :

```
CAGCTGGAGACCAAACCTCTAGGATCTAGGTACCCCCCTGGCAGCAGAGCTGATTTATGGTGCCCTGACAGCCTAATATCACCAAATTACCCACACAAT
GGAGAGAGGCTGGGCTGCCTCCTGGCAGTCGCATAAAGGACGCTGAGCCTGGGGCAGCCCTGCCGCCTTCTGATCCAGACCTCCTTTGGGTCTGTGAGAGA
GGGGCTGGTGCAGACAGAGGCACACCCAGGCATGCCATCATCAACTGGGGTATAAGTAGGAAGCCTTGGGGACAGCTCTGTACCCTGCACTCGGGAGCC
AGTAGCACTCGGACGACTACTTCTGTCCCACAAGCTGCAACAGAGCCAGCCATGAATTGCAGACAGTTCTTGTCTCCCACTGCAGCCGTGACAGCTCTG
GGGTGGTGGGGCAGCAAGATGTTCTCCTGCAACCGCAGCAGCTTTTCAGGGGCTAGTGGAGGAGGGGCCGATTCAGCTCTTCTAGGAGCTTCGGAGGG
GGCAGCTCTGGGGCCTGTGGGAGGGGAGGTGGTGGCAGTTTTGGCTCTAGCTACGGTGGAGGATCTGGGGCGGTTTTAGTGCTGGTAGCTTTGGGGACA
TTCTAGGGCTTCAGTGGTGGTTCTGGAGGAGGCTTTGGCGGTGGATTTGGAGGAGGCTTTGGGGCTTTAGTGGTGGTTCTGGAGGAGGCTTT
```

**Questions**

- 1- Copier-coller cette séquence dans une fenêtre de requête de [BLAST](#). Lancer une requête de recherche de séquence parmi les banques de données (nucleotide collection nr nt) pour les séquences très similaires (megablast).
- 2- Les quatre meilleurs alignements fournis par BLAST correspondent à deux espèces différentes, un carnivore et un primate. Donner le nom de ces espèces.
- 3- D'après le résultat d'alignement, identifier l'espèce à laquelle appartient l'échantillon de sang (nom latin et nom commun). Justifier la réponse.
- 4- D'après la fiche Wikipedia correspondant à l'espèce identifiée, proposer des hypothèses sur l'identité de la bête du Gévaudan.
- 5- Afficher la fiche de renseignement NCBI de la séquence codante donnant le meilleur score d'alignement. En déduire :
 - . le nom du gène auquel appartient cette séquence ;
 - . la protéine codée ;
 - . le rôle de cette protéine



Activité n°2 : construction d'un arbre phylogénétique à l'aide de l'algorithme ClustalW

| | |
|----------------|---|
| Durée | 1h |
| Modalités | Rendre le travail sous forme d'un document word (ou autre traitement de texte) avec saisies écrans justificatives. Penser à mettre son nom dans le nom du fichier afin de le reconnaître rapidement. |
| Liens utilisés | ClustalW2 : http://www.ebi.ac.uk/Tools/msa/clustalw2/ NCBI : http://www.ncbi.nlm.nih.gov/ |

Problématique :

Un arbre phylogénétique est un arbre schématique ou **dendogramme** qui montre **les relations de parentés évolutive** entre des groupes d'êtres vivants. Chacun des nœuds de l'arbre représente l'ancêtre commun de ses descendants. Charles Darwin fut un des premiers scientifiques à proposer une histoire des espèces représentée sous la forme d'un arbre. Cette histoire de l'Evolution des espèces a été confirmée et affinée par les alignements de séquences. Les gènes des **ARN ribosomiaux**, présents dans toutes les espèces ont été largement utilisés pour ces alignements, mais d'autres séquences peuvent aussi être utilisées.

On se propose de réaliser un arbre phylogénétique pour la protéine myoglobine à partir d'un alignement multiple de séquences réalisé par [ClustalW2](#).



Questions

1- Rechercher sur le portail NCBI, la fiche de référence « Genbank » de la myoglobine pour les 7 espèces suivantes :

- . *homo sapiens* (l'homme)
- . *iguana iguana* (l'iguane)
- . *gallus gallus* (le poulet)
- . *pongo abelii* (l'orang outang)
- . *danio rerio* (le poisson zèbre)
- . *rattus norvegicus* (le rat)
- . *equus caballus* (le cheval)

- Aller sur le site NCBI <http://www.ncbi.nlm.nih.gov/>
- Sélectionner « Protein » dans le menu déroulant principal (en haut au centre)
- Requête pour l'homme : « myoglobin homo sapiens »

2- Pour chaque espèce,

Rechercher à l'aide du portail NCBI (Protein) la séquence au format FASTA.

- Sélectionner « FASTA » sous le titre de la fiche Genbank.
- Copier la séquence en acides aminés avec le nom de la séquence « >..... »

Ouvrir ClustalW2.

- Copier-coller cette séquence dans la fenêtre de requête de ClustalW2.

Faire de même pour les six autres séquences.



2- Pour chaque espèce, rechercher à l'aide du portail NCBI (Protein) la séquence au format FASTA.

- Sélectionner « FASTA » sous le titre de la fiche Genbank.
- Copier la séquence en acides aminés avec le nom de la séquence « >..... »

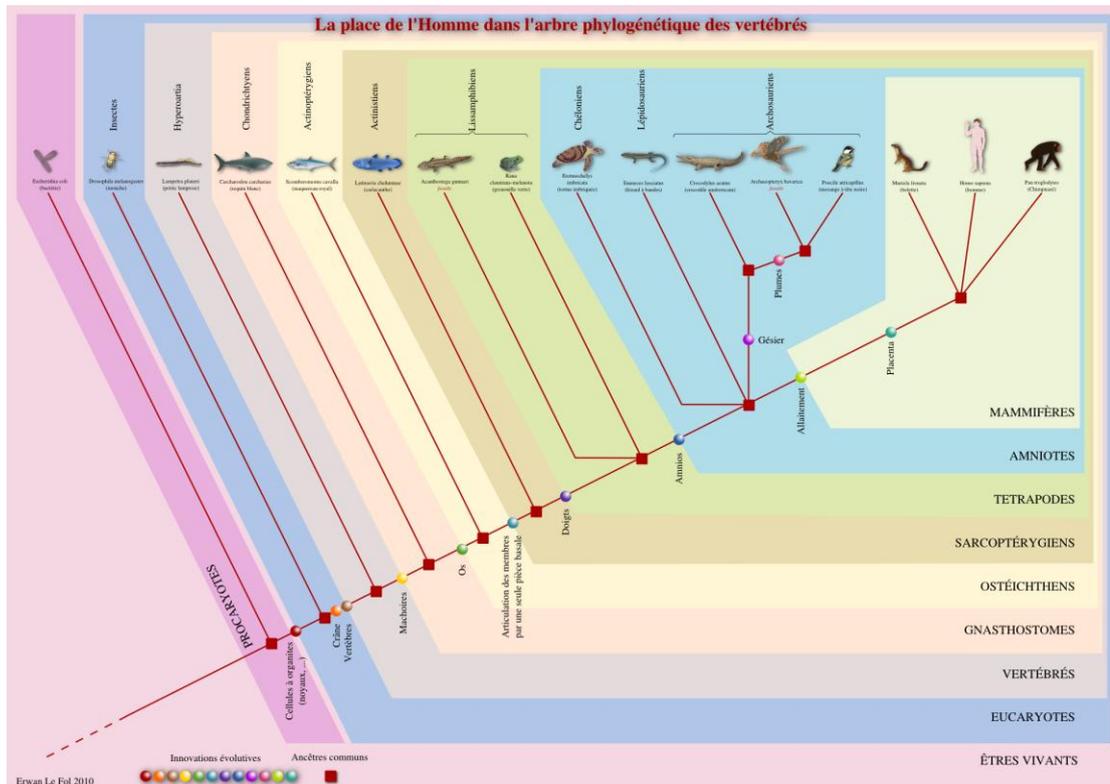
Ouvrir ClustalW2. Copier-coller cette séquence dans la fenêtre de requête de ClustalW2. Faire de même pour les six autres séquences.

3- Exécuter ClustalW2 sur les 7 séquences saisies (« SUBMIT »). Interpréter le résultat de l'alignement multiple (demander un graphe avec les couleurs symboliques des acides aminés).

Attention : pour que l'arbre phylogénétique soit facilement lisible, **il ne faut garder dans le nom de la séquence que le nom de l'espèce.**

4- Interpréter le résultat de l'alignement multiple (demander un graphe avec les couleurs symboliques des acides aminés).

5- Demander la réalisation de l'arbre phylogénétique correspondant. Interpréter cet arbre en le comparant à l'arbre phylogénétique des vertébrés construit à partir des seuls critères anatomiques (à télécharger).



Votre avis nous intéresse, merci de répondre à notre enquête concernant ce scénario.

Elève, cliquer [ici](#).

Professeur, cliquer [ici](#).