

Corrigé des activités « Alignement de séquences »

ACTIVITE 1 : BLAST et la Bête du Gévaudan... :

Rendre le travail sous forme d'un document word ou autre traitement de texte avec saisies écrans justificatives.

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 :

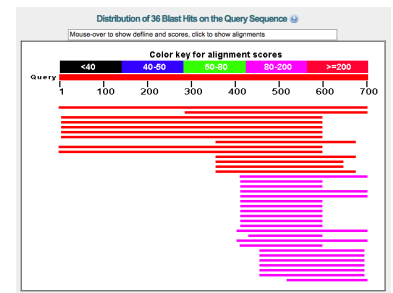
```
CAGCTGGAGACCAAACTCTAGGATCTAGGTCACCCCTTGGCAGCAGAGCTGATTTATGGTGCCCTGACAGCCT
AATATCACCAATTACCCACACAATGGAGAGAGGCTGGGCTGCCTCCTGGCAGTCGCTAAAGGACGCTGAGCCT
GGGGCAGCCCTGCCGCTTCTGATCCCAGACCTCCTTTGGGTCTGTGAGAGAGGGGGCTGGTGCAGACAGAGGCA
CACCCAGGCATGCCATCATCAACTGGGGGTATAAGTAGGAAGCCTTGGGGACAGCTCTGTACCCTGCACTCGGGA
GCCAGTAGCACTCGGACGACTACTTCTGTCCCACAAGCTGCAACAGAGCCAGCCATGAATTGCAGACAGTTCTT
GTCCTCCCCTGACAGCCGTGACAGCTCTGGGGGTGGTGGGGGCAGCAAGATGTTCTCTGCAACCGCAGCAGCTT
TTCAGGGGCTAGTGGAGGAGGGGGCCGATTGAGCTCTTCTAGGAGCTTCGGAGGGGGCAGCTCTGGGGCCTGTGG
GAGGGGAGGTGGTGGCAGTTTTGGCTCTAGCTACGGTGGAGGATCTGGGGGCGGTTTTAGTGCTGGTAGCTTTGG
GGGACATTCTAGGGGCTTCAGTGGTGGTTCTGGAGGAGGCTTTGGCGGTGGATTTGGAGGAGGCTTTGGGGGCTT
TAGTGGTGGTTCTGGAGGAGGCTTT
```

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).

The screenshot shows the NCBI BLAST search page. The 'Enter Query Sequence' section contains the DNA sequence from the text. A blue arrow points to the 'Query subrange' fields. The 'Choose Search Set' section has 'Database' set to 'Nucleotide collection (nr/nt)' and 'Organism' set to 'Optional'. A blue arrow points to the 'Database' dropdown. The 'Program Selection' section has 'Optimize for' set to 'Highly similar sequences (megablast)'. A blue arrow points to this option. At the bottom, the 'BLAST' button is highlighted with a blue arrow, and the search parameters are summarized: 'Search database Nucleotide collection (nr/nt) using Megablast (Optimize for highly similar sequences)'.

2) Les quatre meilleurs alignements fournis par BLAST correspondent à deux espèces différentes, un carnivore et un primate. Indiquer ces espèces.



Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenBank Graphics Distance tree of results

	Description	Max score	Total score	Query cover	E value	Ident	Accession
☐	Canis familiaris keratin (KRT9) gene, complete cds	1293	1293	100%	0.0	100%	AF000949.1
☐	Canis lupus familiaris keratin 9 (KRT9), mRNA	765	765	59%	0.0	100%	NM_001014285.1
☐	Homo sapiens FOSMID clone ABC12-46333900B21 from chromosome 17, complete sequence	427	427	84%	6e-116	80%	AC216692.2
☐	Homo sapiens keratin 9 (KRT9), RefSeqGene on chromosome 17	427	427	84%	6e-116	80%	NG_008300.1
☐	Homo sapiens chromosome 17, clone RP13-415G19, complete sequence	427	427	84%	6e-116	80%	AC130686.15
☐	Homo sapiens chromosome 17, clone RP11-156A24, complete sequence	427	427	84%	6e-116	80%	AC019349.28
☐	Homo sapiens chromosome 17, clone CTD-3226O14, complete sequence	427	427	84%	6e-116	80%	AC105208.15
☐	PREDICTED: Alluopoda melanoleuca keratin, type I cytoskeletal 9-like (LOC100476730), mRNA	390	390	45%	8e-105	89%	XM_002922160.1
☐	Pan troglodytes BAC clone CH251-417E19 from chromosome 17, complete sequence	383	383	85%	1e-102	79%	AC185986.3
☐	Pan troglodytes chromosome 17 clone CH251-433E19, complete sequence	377	377	85%	6e-101	78%	AC254816.1

Carnivore= *Canis familiaris* (chien)
Primate = *Homo sapiens* (humain)

3) D'après le résultat d'alignement, identifier l'espèce à laquelle appartient l'échantillon de sang (nom latin et nom commun). Justifier votre réponse. Conclure sur la nature de la « bête ».

Comparaison des alignement pour la kératine du chien et celle de l'homme :

Canis familiaris keratin (KRT9) gene, complete cds
Sequence ID: [gb|AF000949.1|AF000949](#) Length: 7135 Number of Matches: 1

Range 1: 1 to 700	GenBank	Graphics	Next Match	Previous Mat
Score	Expect	Identities	Gaps	Strand
1293 bits(700)	0.0	700/700(100%)	0/700(0%)	Plus/Plus
Query 1	CAGCTGGAGACCAAACTCTAGGATCTAGGTCAACCCCTTGGCAGCAGAGCTGATTTATGG	60		
Sbjct 1	CAGCTGGAGACCAAACTCTAGGATCTAGGTCAACCCCTTGGCAGCAGAGCTGATTTATGG	60		
Query 61	TGCCCCCTGACAGCCTAAATATACCAAATTAACCCACACAATGGAGAGAGGCTGGGCTGCC	120		
Sbjct 61	TGCCCCCTGACAGCCTAAATATACCAAATTAACCCACACAATGGAGAGAGGCTGGGCTGCC	120		
Query 121	TCCTGGCAGCTCTAAAGAGCCTGAGCCTGGGGCAGCCCTGCCCTTCTGATCCCAAGA	180		
Sbjct 121	TCCTGGCAGCTCTAAAGAGCCTGAGCCTGGGGCAGCCCTGCCCTTCTGATCCCAAGA	180		
Query 181	CCCTCCTTTGGGCTCTGTGAGAGAGGGGGCTGGTCCAGACAGAGGCACACCCAGGCATGCCA	240		
Sbjct 181	CCCTCCTTTGGGCTCTGTGAGAGAGGGGGCTGGTCCAGACAGAGGCACACCCAGGCATGCCA	240		
Query 241	TCATCAACTGGGGGTATAAGTAGGAAGCCTTGGGGACAGCTCTGACCTGCACATCGGGA	300		
Sbjct 241	TCATCAACTGGGGGTATAAGTAGGAAGCCTTGGGGACAGCTCTGACCTGCACATCGGGA	300		
Query 301	GCCAGTAGCACTCGGACAGCTACTTCTGTCCCAAGCTGCAACAGAGCCAGCCATGAA	360		
Sbjct 301	GCCAGTAGCACTCGGACAGCTACTTCTGTCCCAAGCTGCAACAGAGCCAGCCATGAA	360		
Query 361	TTGCAGACAGTTCTTGTCTCCCACTGACGCTGACAGCTCTGGGGGTGGTGGGGCAG	420		
Sbjct 361	TTGCAGACAGTTCTTGTCTCCCACTGACGCTGACAGCTCTGGGGGTGGTGGGGCAG	420		
Query 421	CAAGATGTTCTCTGCAACCGCAGCAGCTTTTCAAGGGCTAGTGAGAGAGGGGGCCGATT	480		
Sbjct 421	CAAGATGTTCTCTGCAACCGCAGCAGCTTTTCAAGGGCTAGTGAGAGAGGGGGCCGATT	480		
Query 481	CAGCTCTTCTAGAGAGCTCTGGAGGGGGCAGCTCTGGGGCTTGTGGAGGGGAGTGGTGG	540		
Sbjct 481	CAGCTCTTCTAGAGAGCTCTGGAGGGGGCAGCTCTGGGGCTTGTGGAGGGGAGTGGTGG	540		
Query 541	CAGTTTGGCTCTAGCTACGGTGGAGAGATCTGGGGCGGTTTATGCTCTGGAGCTTTGG	600		
Sbjct 541	CAGTTTGGCTCTAGCTACGGTGGAGAGATCTGGGGCGGTTTATGCTCTGGAGCTTTGG	600		
Query 601	GGGACATTCAGGGGCTCTAGCTGTGGTCTCGAGAGAGCTTTGGCGTGGATTGGAGG	660		
Sbjct 601	GGGACATTCAGGGGCTCTAGCTGTGGTCTCGAGAGAGCTTTGGCGTGGATTGGAGG	660		
Query 661	AGGCTTTGGGGGCTTTAGTGGTGGTCTCGAGAGAGCTTT	700		
Sbjct 661	AGGCTTTGGGGGCTTTAGTGGTGGTCTCGAGAGAGCTTT	700		

Homo sapiens keratin 9 (KRT9), RefSeqGene on chromosome 17
Sequence ID: [ref|NG_008300.1|](#) Length: 13218 Number of Matches: 1

Range 1: 4722 to 5323	GenBank	Graphics	Next Match	Previous Mat
Score	Expect	Identities	Gaps	Strand
427 bits(231)	6e-116	487/609(80%)	24/609(3%)	Plus/Plus
Query 7	GAGACCAAACTCTAGGATCTAGGTCAACCCCTTGGCAGCAGAGCTGATTTATGGTGGCCC	66		
Sbjct 4722	GAGACCAAACTCTAGGATCTAGGTCAACCCCTTGGCAGCAGAGCTGATTTATGGTGGCCC	4781		
Query 67	TGACAGCCTAATATACCAAAATTAACCCACACAATGGAGAGAGGCTGGGCTGGCTCTCG	126		
Sbjct 4782	TGACAGCCTAATATACCAAAATTAACCCACACAATGGAGAGAGGCTGGGCTGGCTCTCG	4841		
Query 127	CAGTCGCTAAAGAGCCTGAGCCTGGGGCAGCCCTGGCCCTTCTGATCCAGAGCTCTCT	186		
Sbjct 4842	CAGTCGCTAAAGAGCCTGAGCCTGGGGCAGCCCTGGCCCTTCTGATCCAGAGCTCTCT	4901		
Query 187	TTGGGCTCTGAGAGAGGGGGCTGGTGGAG--ACAGAGGACAGAGGACAGCCATCATC	245		
Sbjct 4902	TTGGGCTCTGAGAGAGGGGGCTGGTGGAG--ACAGAGGACAGAGGACAGCCATCATC	4959		
Query 246	AAGTGGGGGTATAAGTAGGAAGCCTTGGGGACAGCTCTGTATCCCTGCATCGGAGCAG	305		
Sbjct 4960	AAGTGGGGGTATAAGTAGGAAGCCTTGGGGACAGCTCTGTATCCCTGCATCGGAGCAG	5019		
Query 306	TAGCACTCGGACAGCTACTTCTGTCCCAAGCTGCAACAGAGC--CAGCATGAATTGC	364		
Sbjct 5020	TAGCACTCGGACAGCTACTTCTGTCCCAAGCTGCAACAGAGC--CAGCATGAATTGC	5075		
Query 365	AGACAGTCTT--T-GTCT--C-CA-CTGAGC--C---GTGACAGCTCTGG---GGGT	409		
Sbjct 5075	AGACAGTCTTCTCTCTCTCTAGTGGAGCTTCGAGAGGGGAGCTCTGGGGCTCTGGGAG	5135		
Query 410	GGTGGGGGACAGAGATGCTCTCTGCAACCGCAGCAGCTTTTCAAGGGCTAGTGGAGGA	469		
Sbjct 5136	GGTGGGGGACAGAGATGCTCTCTGCAACCGCAGCAGCTTTTCAAGGGCTAGTGGAGGA	5195		
Query 470	GGGGGCCGATTCAGCTCTTCTAGGAGCTTCGAGAGGGGAGCTCTGGGGCTCTGGGAG	529		
Sbjct 5196	GGGGGCCGATTCAGCTCTTCTAGGAGCTTCGAGAGGGGAGCTCTGGGGCTCTGGGAG	5255		
Query 530	GGAGTGGTGGCAGTTTGGCTAGCTACGGTGGAGGATCTGGGGCGGTTTATGCTCT	589		
Sbjct 5256	GGAGTGGTGGCAGTTTGGCTAGCTACGGTGGAGGATCTGGGGCGGTTTATGCTCT	5315		
Query 590	GGTAGCTTT	598		
Sbjct 5316	AGTAG--TTT	5323		

100% d'identité pour la kératine de chien, 80% pour celle de l'homme. La bête est vraisemblablement un *Canis lupus familiaris* ou un *Canis lupus*.

4) D'après la fiche Wikipedia correspondant à l'espèce identifier, proposer des hypothèses sur l'identité de la bête du Gévaudan.

Fiche *Canis lupus familiaris* de Wikipedia :

Le **chien** (*Canis lupus familiaris*) est la sous-espèce domestique de *Canis lupus*, un mammifère de la famille des canidés qui comprend également le **loup gris**, ancêtre sauvage du chien, et le dingo, chien domestique redevenu sauvage.

Le chien est la première espèce animale à avoir été domestiquée par l'Homme pour l'usage de la chasse dans une société humaine paléolithique qui ne maîtrise alors ni l'agriculture ni l'élevage. La lignée du chien se différencie génétiquement de celle du loup gris il y a environ 100 000 ans, et les plus anciens restes confirmés de chien domestique sont vieux de 33 000 ans 2,3 soit plusieurs dizaines de milliers d'années avant toute autre espèce domestique connue.

➔ la bête peut être un loup ou un chien...

5) Afficher la fiche de renseignement NCBI de la séquence codante donnant le meilleur score d'alignement. A partir de cette fiche et des liens qui lui sont associés, déduire :

. le nom du gène auquel appartient cette séquence ;

Gène de la Kératine, KRT9

. la protéine codée ;

La Kératine

. le rôle de cette protéine.

Extrait de la fiche « Kératine » de Wikipedia :

La kératine est une protéine, synthétisée et utilisée par de nombreux êtres vivants comme élément de structure, et également l'exemple-type de protéine fibreuse. C'est le constituant principal des phanères (poils, plumes, cornes, ongles, becs de nombreux animaux) (...)

La kératine est insoluble, et sa présence dans l'épiderme de certains animaux, notamment les mammifères, leur garantit une peau imperméable. (...)

La molécule de kératine est hélicoïdale et fibreuse ; elle s'enroule autour d'autres molécules de kératine pour former des filaments intermédiaires. Ces protéines contiennent un haut taux d'acides aminés à base de soufre, principalement la cystéine, qui forment des ponts disulfure entre les molécules, conférant sa rigidité à l'ensemble. (...)

Chez l'être humain, la kératine est fabriquée par les kératinocytes, cellules se trouvant dans la couche profonde de l'épiderme.

ACTIVITE 2 : construction d'un arbre phylogénétique à l'aide de l'algorithme ClustalW2:

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.

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 ribosomaux**, 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 des 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 »

The top screenshot shows the NCBI homepage. The search bar is set to 'Protein' and the query is 'myoglobin homo sapiens'. The search button is labeled 'Search'. The bottom screenshot shows the search results page. The search bar is still 'Protein' and the query is 'myoglobin homo sapiens'. The search button is labeled 'Search'. The results are displayed as a list of three entries, all for the 154 aa protein. The first entry is 'myoglobin [Homo sapiens]' with accession number NP_976312.1. The second entry is 'myoglobin [Homo sapiens]' with accession number NP_976311.1. The third entry is 'myoglobin [Homo sapiens]' with accession number NP_005359.1. The results are sorted by Default order and displayed 20 per page.

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.

[Display Settings:](#) ☒ GenPept

myoglobin [Homo sapiens]

NCBI Reference Sequence: NP_976312.1

[FASTA](#) [Graphics](#)

[Go to:](#) ☒

```

LOCUS      NP_976312                154 aa          linear   PRI 26-JAN-2014
DEFINITION myoglobin [Homo sapiens].
ACCESSION  NP_976312
VERSION    NP_976312.1  GI:44955888
DBSOURCE   REFSEQ: accession NM\_203378.1
KEYWORDS   RefSeq.
SOURCE     Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
            Catarrhini; Hominidae; Homo.
  
```

- Copier la séquence en acides aminés avec le nom de la séquence « >..... »

[Display Settings:](#) ☒ FASTA

myoglobin [Homo sapiens]

NCBI Reference Sequence: NP_976312.1

[GenPept](#) [Graphics](#)

```

>gi|44955888|ref|NP_976312.1| myoglobin [Homo sapiens]
MGLSDGEWQLVLNVWGKVEADIPGHGQEV LIRLFKGH PETLEKFDKFKHLKSEDEM KASEDLKKHGATVL
TALGGILKKKGHHEAEIKPLAQSHATKHKIPVKYLEFISECIIQVLQSKHPGDFGADAQGAMNKALELFR
KDMASNYKELGFQG
  
```

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

Faire de même pour les six autres séquences.

```

> [Homo sapiens]
MGLSDGEWQLVLNVWGKVEADIPGHGQEV LIRLFKGH PETLEKFDKFKHLKSEDEM KASEDLKKHGATVL TALG
GILKKKGHHEAEIKPLAQSHATKHKIPVKYLEFISECIIQVLQSKHPGDFGADAQGAMNKALELFR KDMASNYKELG
FQG
  
```

```

> [Iguana iguana]
MGLSDQEWQKVIDIWGKVEPELPAHGQEV IIRLFQKHPETQEKFDFKHLKSLDEMKSSEEIKKHGTTVLTALGKIL
KQKGHHDAELAPLAQSHATKHKIPVKYLEFICEVIVGVIAEKHSADFGSESQAAMRKALEVFRNDMASKYKEHGFG
G
  
```

```
>[Gallus gallus]
MGLSDQEWQVLTIVWGKVEADIAGHGHEVLMRLFDHPETLDRFDKFKGLKTPDQMKGSSEDLKKHGATVLTQLG
KILKQKGNHESELKPLAQTHATKHKIPVKYLEFISEVIIKVIAEKHAADFGADSQAAMKKALELFRNDMASKYKEFGF
QG

>[Pongo abelii]
MGLSDGEWQVLVNVWGKVEADIPSHGQEVILIRLFKGHPETLEKFDKFKHLKSEDEMKASEDLKKHGATVLTALG
GILKKKGHHAEIKPLAQSHATKHKIPVKYLEFISESIIQVLQSKHPGDFGADAQGAMNKALELFRKDMASNYKELG
FQG

> [Danio rerio]
MADHDLVLKCVGWAVEADYAANGGEVLNRLFKEYPDTLKLFPKFSGISQGDLAGSPAHAHGATVLTAKLKGELLKAK
GDHAALLKPLANTHANIHKVALNNFRLITEVLVKVMAEKAGLDAAGQGALRRVMDAVIDGIDIDGYKKEIGFAG

>[Rattus norvegicus]
MGLSDGEWQMVNIWVGKVEGDLAGHGQEVILISLFAHPETLEKFDKFKNLKSEEMKSSSEDLKKHGCTVLTALG
TILKKKGQHA AEIQPLAQSHATKHKIPVKYLEFISEVIIQVLKKRYS GDFGADAQGAMSKALELFRNDIAAKYKELG
FQG

>[Equus caballus]
MGLSDGEWQVNLVNVWGKVEADIAGHGQEVILIRLFTGHPETLEKFDKFKHLKTEAEMKASEDLKKHGTVVLTALG
GILKKKGHHAEELKPLAQSHATKHKIPVKYLEFISDAIIHVLHSHKHPGDFGADAQGAMTKALELFRNDIAAKYKELG
FQG
```

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.

Results for job clustalw2-l20140129-161941-0030-59604016-oy

Alignments Result Summary Guide Tree Phylogenetic Tree Submission Details

Download Alignment File Hide Colors Send to ClustalW2_Phylogeny

CLUSTAL 2.1 multiple sequence alignment

```
[Homo      MGLSDGEWQVLVNVWGKVEADIPGHGQEVILIRLFKGHPETLEKFDKFKHLKSEDEMKASE 60
[Pongo     MGLSDGEWQVLVNVWGKVEADIPSHGQEVILIRLFKGHPETLEKFDKFKHLKSEDEMKASE 60
[Equus     MGLSDGEWQVNLVNVWGKVEADIAGHGQEVILIRLFTGHPETLEKFDKFKHLKTEAEMKASE 60
[Rattus    MGLSDGEWQMVNIWVGKVEGDLAGHGQEVILISLFAHPETLEKFDKFKNLKSEEMKSSSE 60
[Iguana    MGLSDGEWQKVIDIWGKVEPELPAHGQEVILIRLFKGHPETLEKFDKFKHLKSLDEMKSSSE 60
[Gallus    MGLSDGEWQVLTIVWGKVEADIAGHGHEVLMRLFDHPETLDRFDKFKGLKTPDQMKGSE 60
[Danio     --MADH--DLVLKCVGWAVEADYAANGGEVLNRLFKEYPDTLKLFPKFSGIS--QGDLAGSP 55
          ::*  : *  ** ** : ..** **  :*: . * ** . :  : *
```

```
[Homo      DLKKHGATVLTALGGILKKKGHHEAEIKPLAQSHATKHKIPVKYLEFISECIIQVLQSKH 120
[Pongo     DLKKHGATVLTALGGILKKKGHHEAEIKPLAQSHATKHKIPVKYLEFISESIIQVLQSKH 120
[Equus     DLKKHGTVLTALGGILKKKGHHEAELKPLAQSHATKHKIPVKYLEFISDAIIHVLHSHK 120
[Rattus    DLKKHGCTVLTALGTILKKKGQHA AEIQPLAQSHATKHKIPVKYLEFISEVIIQVLKKRY 120
[Iguana    EIKKHGTTVLTALGKILKQKGHDAELAPLAQSHATKHKIPVKYLEFICEVIVGVIAEKH 120
[Gallus    DLKKHGATVLTQLGKILKQKGNHESELKPLAQTHATKHKIPVKYLEFISEVIIKVIAEKH 120
[Danio     AVAAHGATVLTAKLGEELKAGDHAALLKPLANTHANIHKVALNNFRLITEVLVKVMAEKA 115
          :  ** .** . ** : ** ** : : ***:***. ** : : : : ** : *
```

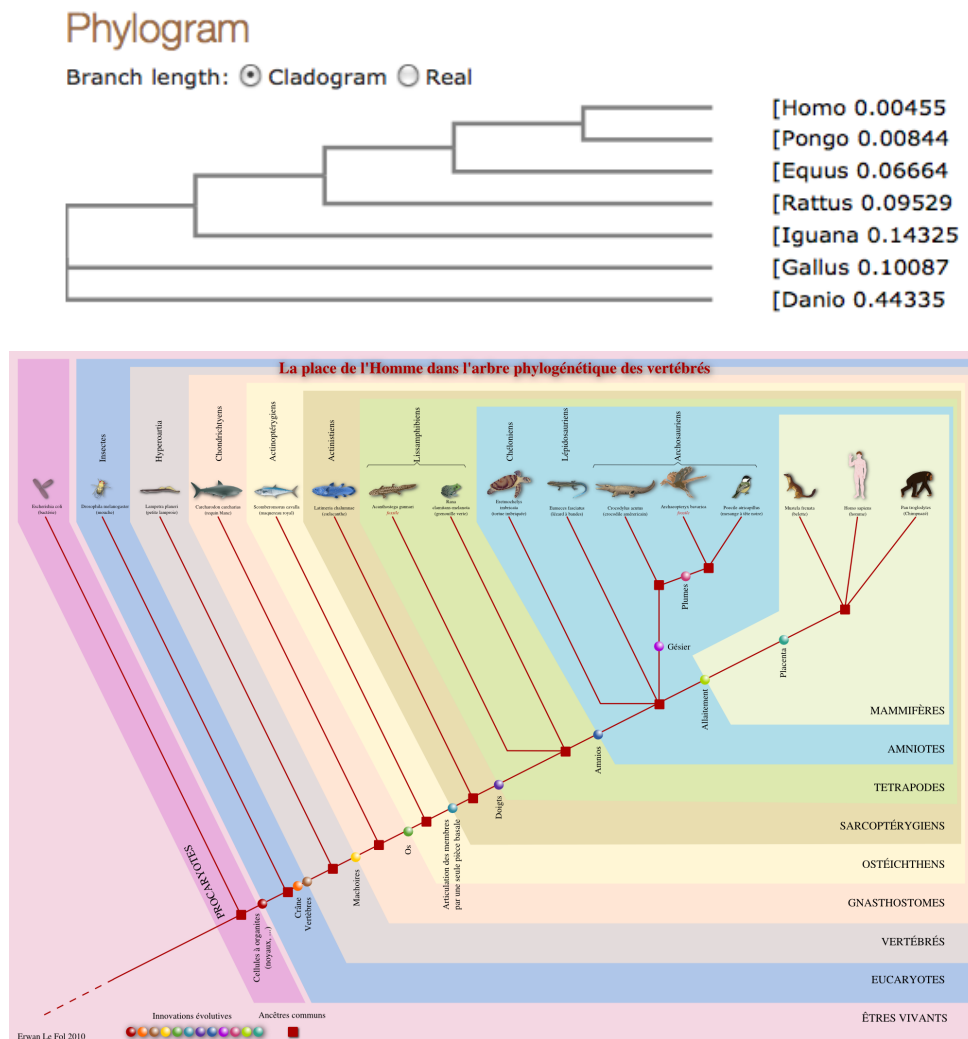
```
[Homo      PGDFGADAQGAMNKALELFRKDMASNYKELGFQG 154
[Pongo     PGDFGADAQGAMNKALELFRKDMASNYKELGFQG 154
[Equus     PGDFGADAQGAMTKALELFRNDIAAKYKELGFQG 154
[Rattus    SGDFGADAQGAMSKALELFRNDIAAKYKELGFQG 154
[Iguana    SADFGSESQAAMRKALEVFRNDMASKYKEHGFQG 154
[Gallus    AADFGADSQAAMKKALELFRNDMASKYKEFGFQG 154
[Danio     G--LDAAGQCALRRVMDAVIDIDIDGYKKEIGFAG 147
          :.: .*. : : : : . * : . *** ** *
```

La myoglobine contient 154 aa sauf pour le poisson (147).

Certains aa sont conservés chez toutes les espèces, ils sont marqués par une *.

Parmi les aa modifiés le code couleur permet de montrer que certains aa sont substitués par un aa présentant des propriétés voisines (ex : aa apolaires en rouge) ; d'autres sont remplacés par un aa avec d'autres propriétés (couleur différente).

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).



L'arbre établi par alignement de séquence de la protéine Myoglobine correspond à celui établi par les critères anatomiques.

Par ordre de proximité évolutive décroissante :
Homme, Orang-Outang, Cheval, Rat, Iguane, Poulet, Poisson.