

Construisez un arbre phylogénétique !



Il est possible de construire un arbre phylogénétique à partir de différents types de données:

- les **données morphologiques** (écailles ou plumes, présence de certains os du crâne, température corporelle, ...). Il existe quelques centaines de caractères définis dans ce but par les spécialistes.

- les **données moléculaires** (séquences d'ADN ou de protéines). Des mutations modifient les séquences de l'ADN et par conséquent des protéines au cours de l'évolution.

...grâce à PhiloPhylo, un programme bioinformatique, vous allez construire des arbres phylogénétiques obtenus sur la base de séquences de protéines et découvrirez que les protéines ont aussi la mémoire de leurs ancêtres...



Choisissez une protéine

- Insuline** (permet au sucre d'entrer dans les cellules)
- TPIS** (impliquée dans le métabolisme des sucres)
- MetRS** (impliquée dans la synthèse des protéines)
- Histone H4** (importante pour la structure des chromosomes)
- ATPA** (impliquée dans la production d'énergie)
- Hormone de croissance** (fait grandir les enfants...)
- Cytochrome B** (impliquée dans la production d'énergie)



Dans la base de données

recherchez les séquences de cette protéine dans différentes espèces

OK

Joyeuses découvertes !



Choisissez une protéine



L'idéal est de choisir une protéine qui vous intéresse et qui est connue dans les espèces que vous voulez comparer... ce n'est pas toujours le cas:

- Les biologistes savent qu'une protéine existe dans une espèce donnée, mais elle n'a pas encore été caractérisée: par exemple, les giraffes ont de l'insuline, mais sa séquence n'est pas encore connue.
- la protéine n'existe pas dans une espèce: les bactéries, les levures ou le maïs par exemple, n'ont pas d'insuline. D'autres protéines, comme le cytochrome B, sont appréciées des phylogénéticiens, car leurs séquences sont connues dans de très nombreuses espèces, y compris dans des espèces disparues comme le dodo ou le mammouth !



...et recherchez les séquences de cette protéine chez différentes espèces... autrement dit, allez pêcher l'équivalent de votre protéine (humaine) dans d'autres espèces comme la grenouille, le bœuf ou le lapin.

Pour cela, il faut utiliser un programme bioinformatique (appelé 'Blast') qui va 'pêcher' dans une base de données toutes les protéines qui ressemblent à la protéine de départ. La base de données que nous utilisons est appelée Swiss-Prot. Elle contient des informations sur ~230'000 protéines provenant de ~10'000 espèces différentes.



Comparez les séquences des protéines... dans le jargon bioinformatique, on parle d'alignement de séquences.

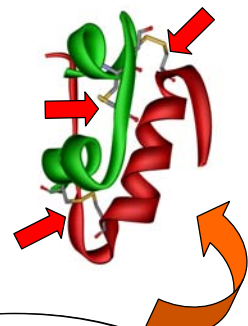
Pour comprendre cette étape, il faut se rappeler que les protéines peuvent être comparées à des colliers de longueur variable, composés de 20 perles différentes: les acides aminés. Les acides aminés sont représentés par des lettres majuscules (G, E, N, I, E...). Construire un alignement consiste à mettre en relation les colliers de perles les uns avec les autres en fonction de la similarité de leurs perles (comparer de façon bio-logique les différents colliers de perles). Au cours de l'évolution, les acides aminés qui jouent un rôle important pour la structure ou la fonction d'une protéine ne sont pas ou peu modifiés et se retrouvent à la même place dans toutes les espèces. Les acides aminés moins 'vitaux' pour la fonction de la protéine peuvent être plus ou moins modifiés en fonction du moment auquel les espèces ont divergé.



Voici un alignement des séquences de l'insuline de différentes espèces:

Grenouille	MALWMQCLELVLVFFSTPNT-EALVNHLCSSHLVEALYLVCSDRFFFYVVKDMLQAL-VSGPQD---NELDGMQLQPEYQKMRGIVECCCHSTLFLQIESYCN
Poulet	MALWIRSLFLALVFGSPGTSYAAANHLCCSHLVEALYLVCGERGFFYSNARRDVKQL-VSSPLR---GEAGVLPFQEEYKVKRGIVECCCHNTSLYQIENYCN
Chimpanzé	MALWMRLLFLVLLALWGPDPASAFVNHLCSSHLVEALYLVCGERGFFYFTRREADLQ-VGQVELGGGPGAGSLQPLALEGSLQKRGIVECCCTSIICSLYQIENYCN
Homme	MALWMRLLFLVLLALWGPDPAAAFVNHLCSSHLVEALYLVCGERGFFYFTRREADLQ-VGQVELGGGPGAGSLQPLALEGSLQKRGIVECCCTSIICSLYQIENYCN
Orang-outan	MALWMRLLFLVLLALWGPDPAAAFVNHLCSSHLVEALYLVCGERGFFYFTRREADLQ-VGQVELGGGPGAGSLQPLALEGSLQKRGIVECCCTSIICSLYQIENYCN
Rat	MALWMRFLFLVLLALVWEPKPAQAFVNHLCSSHLVEALYLVCGERGFFYFTRSRREVDPQ-VPQLELGGGPEAGDLQTLALEVARQKRGIVDCCCTSIICSLYQIENYCN
Souris	MALLVHFLLFLVLLALWEPKPTQAFVNHLCSSHLVEALYLVCGERGFFYFTRSRREVDPQ-VQLELGGSP--GDQLTALAEVARQKRGIVDCCCTSIICSLYQIENYCN
Bœuf	MALWTRKFLFLVLLALWVPPPARAFVNHLCSSHLVEALYLVCGERGFFYFTRSRREVDPQ-VGALLEGSPGAGGLE---GPPCKRGIVECCASVHSLYQIENYCN
Poisson	MAYWLQAGALVLLVYVSSVSTNPGTF--HLCCSHLVDALYLVCSPTEFENEL--DVEPLLGLVPPKSAQETEVADFAFKDHAELIRKRGIVECCCHKPTEFELQIENYCN

■ acides aminés identiques.
■ acides aminés avec les mêmes propriétés biochimiques.
■ acides aminés peu similaire du point de vue biochimique.
■ acides aminés non similaire du point de vue biochimique.

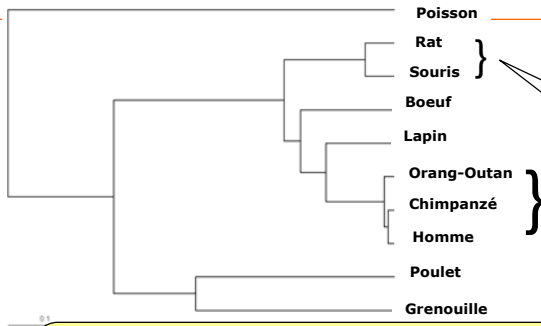


Construisez un arbre phylogénétique...

Pour construire un arbre, il faut utiliser des programmes bioinformatiques qui évaluent les différences ou similitudes observées dans l'alignement. Les résultats peuvent ensuite être visualisés à l'aide d'un arbre phylogénétique. Plus le nombre d'acides aminés est différent entre les protéines, plus les espèces seront 'éloignées' dans les différents embranchement de l'arbre et plus les espèces seront considérées comme 'éloignées' dans l'évolution. Les embranchements (ou noeuds) correspondent à des organismes ancestraux hypothétiques.

Les 6 'C' sont conservés au cours de l'évolution, parce qu'ils sont importants pour la structure de l'insuline: ils forment 3 'ponts'.

Voici un arbre phylogénétique construit à partir de l'alignement des séquences d'insuline:



Cet arbre confirme que les singes et l'homme sont proches dans l'évolution (issus d'un ancêtre commun), de même les souris et les rats, mais que ces derniers sont relativement 'éloignés' du poisson, de la grenouille ou du poulet.

- Selon la protéine ou les caractères morphologiques choisis, les programmes utilisés, les arbres phylogénétiques peuvent être différents... ne vous inquiétez pas... cela donne parfois du fil à retordre même aux spécialistes !
- Les résultats de ces analyses 'moléculaires' sont importants pour les biologistes parce qu'ils sont complémentaires aux analyses des caractères morphologiques: ils ont permis de confirmer beaucoup d'hypothèses sur l'évolution de la vie.
- Il est maintenant possible de construire un arbre en tenant compte à la fois des données morphologiques et des données moléculaires.

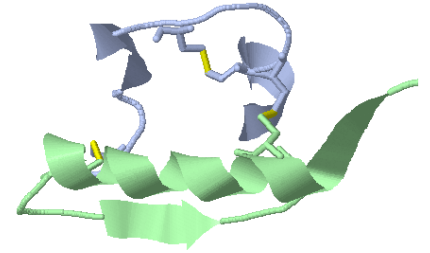
Arbre phylogénétique (insuline)

L'insuline est une hormone cruciale pour le métabolisme du sucre. Elle est fabriquée par le pancréas puis secrétée dans la circulation sanguine. Elle régule le taux de sucre dans le sang en permettant au sucre d'entrer dans les cellules (et donc les tissus). En cas de manque d'insuline (comme c'est le cas dans certaines formes de diabète), il est nécessaire de faire des injections d'insuline.

Séquence en acides aminés de l'insuline humaine

```
MALWMRLPLLLALLALWGPDPAAAFVNQHLGCSHLVEALYLVC
GERGFFYTPKTRREAEDLQVGVQVELGGGPGAGSLQPLALEGSL
QKRGIIVEQCCTSICSLYQLENYCN
```

Structure 3D de l'insuline humaine



Alignement multiple des séquences des insulines des différentes espèces

Espèce	M	A	L	W	M	R	L	P	L	L	A	L	A	L	W	G	P	D	P	A	A	F	V	N	Q	H	L	G	S	H	L	V	E	A	L	Y	L	V	C		
Grenouille	M	A	L	W	M	R	L	P	L	L	A	L	A	L	W	G	P	D	P	A	A	F	V	N	Q	H	L	G	S	H	L	V	E	A	L	Y	L	V	C		
Poulet	M	A	L	W	I	R	S	L	P	L	L	A	L	V	F	S	G	P	G	T	S	Y	A	A	N	H	L	G	S	H	L	V	E	A	L	Y	L	V	C		
Chimpanzé	M	A	L	W	M	R	L	L	P	L	L	A	L	A	L	W	G	P	D	P	A	A	F	V	N	Q	H	L	G	S	H	L	V	E	A	L	Y	L	V	C	
Homme	M	A	L	W	M	R	L	L	P	L	L	A	L	A	L	W	G	P	D	P	A	A	F	V	N	Q	H	L	G	S	H	L	V	E	A	L	Y	L	V	C	
Orang-outan	M	A	L	W	M	R	L	L	P	L	L	A	L	A	L	W	G	P	D	P	A	A	F	V	N	Q	H	L	G	S	H	L	V	E	A	L	Y	L	V	C	
Rat	M	A	L	W	M	R	F	L	P	L	L	A	L	V	L	W	E	P	K	P	A	Q	A	F	V	K	H	L	G	S	H	L	V	E	A	L	Y	L	V	C	
Souris	M	A	L	L	V	H	F	L	P	L	L	A	L	A	L	W	E	P	K	P	T	Q	A	F	V	K	H	L	G	S	H	L	V	E	A	L	Y	L	V	C	
Bœuf	M	A	L	W	T	R	L	R	P	L	L	A	L	A	L	W	P	P	P	P	A	R	A	F	V	N	Q	H	L	G	S	H	L	V	E	A	L	Y	L	V	C
Poisson	M	A	V	W	L	Q	A	G	A	L	L	V	L	V	V	S	S	V	T	N	P	G	T	P	H	L	G	S	H	L	V	D	A	L	Y	L	V	C			

les colonnes colorées mettent en évidence les positions conservés entre les séquences des différentes espèces:

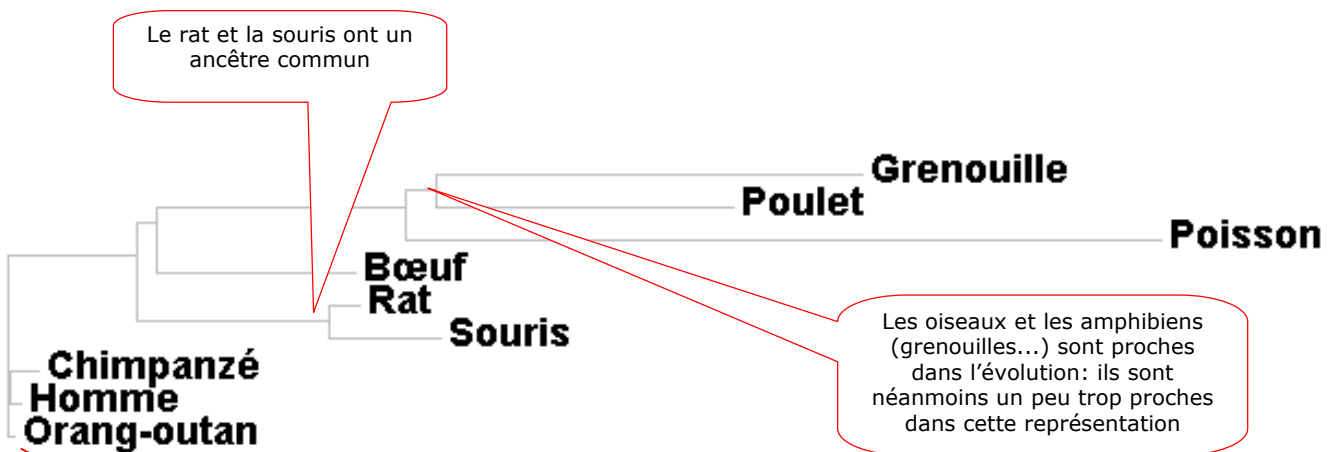
- acides aminés identiques.
- acides aminés avec les mêmes propriétés biochimiques.
- acides aminés peu similaire du point de vue biochimique.
- acides aminés non similaire du point de vue biochimique.

Calcul des différences entre les séquences



Arbre phylogénétique

Représentation graphique des différences et similitudes observées dans l'alignement multiple des séquences des différentes espèces



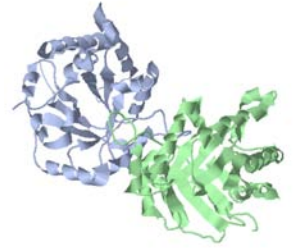
Arbre phylogénétique (TPIS)

TPIS (Triosephosphate isomerase de son petit nom) est une protéine impliquée dans le métabolisme des sucres. Cette protéine est intéressante pour la phylogénie, car elle est universelle, autrement dit, elle existe chez toutes les espèces vivantes, de la bactérie à l'homme.

Séquence en acides aminés de TPIS humaine

MAPSRKFFVGGNWKMNGRKQSLGELIGTLNAAKVPADTEVVCAPPTAYIDFARQKLDPKIA
 VAAQNCYKVTNGAFTGEISPGMIKDCGATWVVLGHSERRHVFGESEDELIGQVAHALAE
 LGVIACIGEKLDEREAGITEKVVFEQTKVIADNVKDWKSVVLAYEPVWAI GTGTATPQQ
 AQEVHEKLRGLKSNVSDAVAQSTRI IYGGSVTGATCKELASQPDVDGFLVGGASLKPEF
 VDIINAKQ

Structure 3D de TPIS



Alignement multiple des séquences de TPIS des différentes espèces (extrait)

<input type="checkbox"/> Chimpanzé	TEVVCAPPTAYIDFARQKLD-PKIAVAAQNCYKVTNGAFTGEISPGMIKDCGATWVVLGHSERRHVFGESEDELIGQVAHALAE
<input type="checkbox"/> Homme	TEVVCAPPTAYIDFARQKLD-PKIAVAAQNCYKVTNGAFTGEISPGMIKDCGATWVVLGHSERRHVFGESEDELIGQVAHALAE
<input type="checkbox"/> Bœuf	TEVVCAPPTAYIDFARQKLD-PKIAVAAQNCYKVTNGAFTGEISPGMIKDLGATWVVLGHSERRHVFGESEDELIGQVAHALAE
<input type="checkbox"/> Souris	TEVVCAPPTAYIDFARQKLD-PKIAVAAQNCYKVTNGAFTGEISPGMIKDLGATWVVLGHSERRHVFGESEDELIGQVSHALAE
<input type="checkbox"/> Rat	TEVVCAPPTAYIDFARQKLD-PKIAVAAQNCYKVTNGAFTGEISPGMIKDLGATWVVLGHSERRHVFGESEDELIGQVNHALSE
<input type="checkbox"/> Poulet	TEVVCAPPTAYIDFARQKLD-PKIAVAAQNCYKVTNGAFTGEISPGMIKDLGATWVVLGHSERRHVFGESEDELIGQVAHALAE
<input type="checkbox"/> Mais	/E/VVSPPTVYVFLPVVKSQLR-QEFHVAQNCWVKKGGAFTGEVSAEMLVNLGVPWVILGHSERRALLGSENEFVGDVAVAYALSQ
<input type="checkbox"/> Riz	/E/VVSPPTVYVFLPVVKSQLR-PEIQVAAQNCWVKKGGAFTGEVSAEMLVNLGVPWVILGHSERRALLGSENEFVGDVAVAYALSQ
<input type="checkbox"/> Mouche	TEVVICPPTAYILMYARNLNP-CEGLLAGQNAVYKVAAGAFTGEISPAMLKDIGADWVILGHSERRAIFGSESDALIAEAHALAE
<input type="checkbox"/> Ver	/D/VVAPPTAYILYAKSKLK-AGVLVAAQNCYKVPKGAFTGEISPAMLKDLGLEWVILGHSERRHVFGESESDALIAEAHALAE
<input type="checkbox"/> Levure	/E/VVICPPTAYILDYSVSLVKKPQVTVGAQNAVYKVAAGAFTGEVNSVDQIKDVGAKWVILGHSERRSYFHEDDKFIADTKKFGALGQ
<input type="checkbox"/> Bactérie	CAVAIAPPEMYIDMAKREAEGSHIMLGAQVVDLNLGSAFTGEVNSVDQIKDVGAKWVILGHSERRSYFHEDDKFIADTKKFGALGQ

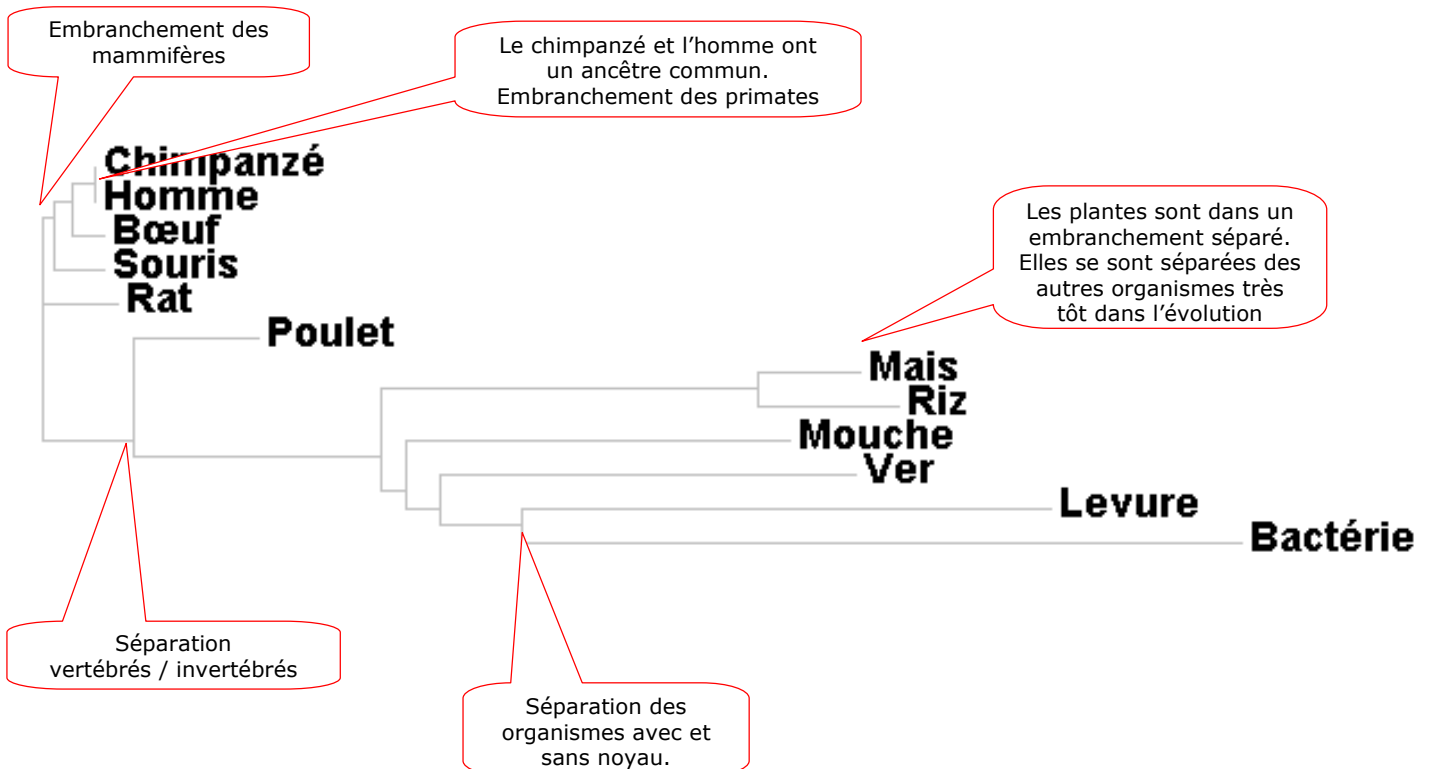
les colonnes colorées mettent en évidence les positions conservées entre les séquences des différentes espèces:
 ■ acides aminés identiques.
 ■ acides aminés avec les mêmes propriétés biochimiques.
 ■ acides aminés peu similaire du point de vue biochimique.
 ■ acides aminés non similaire du point de vue biochimique.

Calcul des différences entre les séquences



Arbre phylogénétique

Représentation graphique des différences et similitudes observées dans l'alignement multiple des séquences des différentes espèces



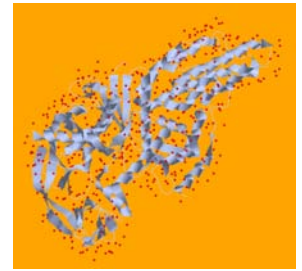
Arbre phylogénétique (MetRS)

MetRS (Methionyl-tRNA synthetase de son petit nom) est une protéine impliquée dans la synthèse des autres protéines. Cette protéine est universelle, car impliquée dans un processus biologique essentiel à la vie.

Séquence en acides aminés de MetRS humaine

```
MELFVSDVPGCLVLAAGRARGRAEVLSTVSPEDCVVPLTRPKVPLVQLDQNYLFTSAICRYVFLSOMEQDULTNQLWEATELQPAALSAALYTLVVGKSGEDVLSVRA
LTHIHSLSRQNCFLAFETSLADIVLHGLVPLQAPLPEELGALHDFUTLSTQPCQRAETVLAQQVLALEPFLQRQPQSPARARAVTNEFEELATLSEELAMAVTAM
EHLSESLPLFQDQVPLPAGESVYLTSAIPYVNYVPELGIIDVLYADYFAYSELEONHLLACTETEVATELAEELQPKQICRTEYLIHADIYEMNLSFTLPSFTTTE
QOKLITQDI PQQLKRFVLDQTVBQLCEBCARFLADRFVEQVCPQCYEERAGQCDKCKGLINAVELEKQCKVCRSCVVOGSHFLDLPKLEELSEMLRSLPQSDMTFRAQF
ITRNLGDLAPFLITELHMDVPLVLESEKVFVWFNATLGLLITANTQDWRNWRNRFQDQVPMADNVFBSLVPFCALGADENTLVHSLIATEYLNIEDGRFSISQV
QVFCMAQDTIIPADLHREFLLIIEEEDGSAFSTGLLENNELANLWYDRAAGNFRFQVQVYVYVWVLTIDQGLLANVLELQRYHGLSEYVSEALSLIETESRNDYIQ
VNEPWRIKGSEADRRAGTVTULAVNIAALLSVGLQPMPTVRSATLQAQLLPPACSLILTNFLCTLFAGHQIGTVSPLFQKLENDQIESLQRFQGGQARTSPKFAVVEVTTAKPQ
QIQALHNDVTRQNIIVRELAQRADREVAEVAKLLDLKQLAVRGRPPAPKRRKKE
```

Structure 3D de MetRS d'une bactérie



Alignement multiple des séquences de MetRS des différentes espèces (extrait)

<input type="checkbox"/> Homme	HLIATEYLNIEDGKFSKSRGVGVFGDMAQDTGIPADIWRFLLYIRPEGQSAFWSWTL LLLKNNSELLNNLGNFINFAGMFVSKFFGG
<input type="checkbox"/> Souris	HI IATEYLNIEDGKFSKSRGIGVFGDMAKDTGIPADIWRFLLYIRPEGQSAFWSWTL LLIKNNSELLNNLGNFINFAGMFVSKFFGG
<input type="checkbox"/> Grenouille	HLVATEYLNIEDGKFSKSRGVGVFGDMAKDTGIPADIWRFLLYVIRPEGQSAFWSWTLMLKNNSELLNNLGNFVNRAGMFVQKFFNG
<input type="checkbox"/> Levure	HLNTEYLNQYENGKFSKSRGVGVFGNNAQDSGISPSVWRYLASVREPESSSHFSDWDFVARNNSELLANLGNFVNRLIKFNKAYNG
<input type="checkbox"/> Riz	TISVTEYLNIEAGKFSKSHGIGVFGNDAKDTNIPPEVWRYL LTNREPEVSTLFTWADLQAKLNSSELLNNLGNFINVLSFVAKPAGA
<input type="checkbox"/> Ver	NLCATEYLNIEDTKFSKSRGTGIFGDAAQGTETIPADIWRFLLYMRPEQDTAFSWDDFVLKVNSELLNNLGNFINRALSFVANSFGG
<input type="checkbox"/> Bactérie	NLFVHGVTVNGAKMSKSRGTFIKASTWLNHFADSLRYYTAKLSSRIDIDLNLLELFVQRVNADIVNKVVNLASFNAGFINKRFDG

les colonnes colorées mettent en évidence les positions conservés entre les séquences des différentes espèces:

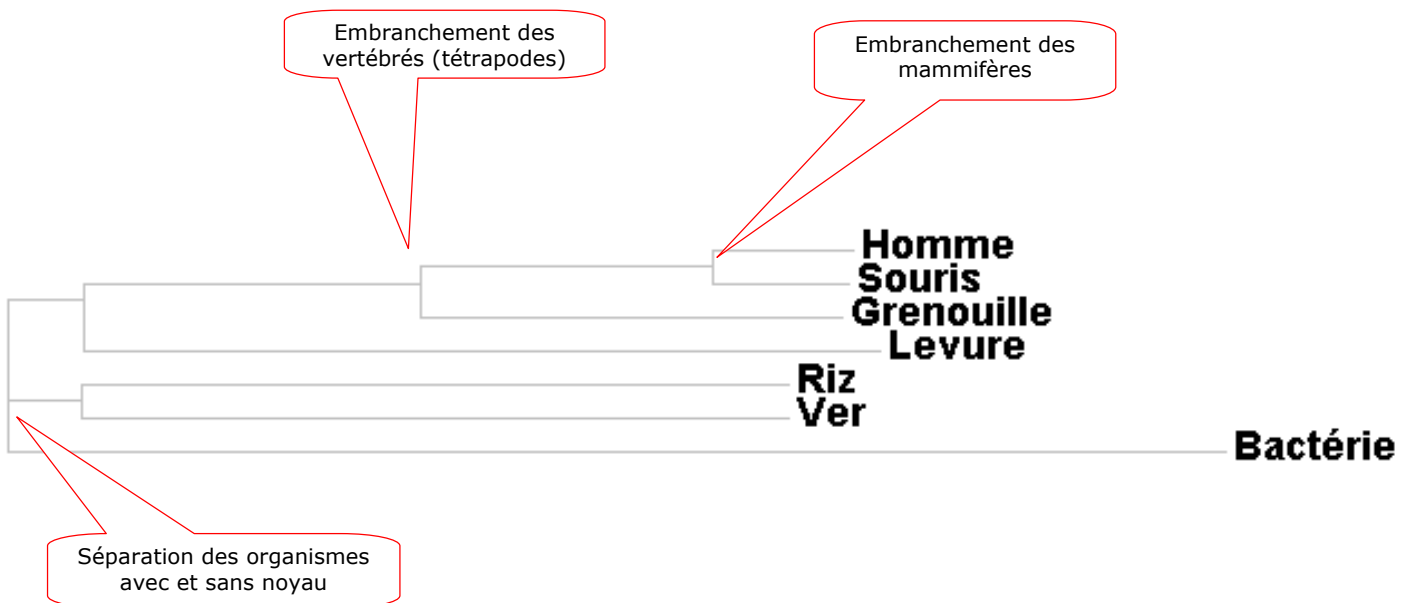
- acides aminés identiques.
- acides aminés avec les mêmes propriétés biochimiques.
- acides aminés peu similaire du point de vue biochimique.
- acides aminés non similaire du point de vue biochimique.

Calcul des différences entre les séquences



Arbre phylogénétique

Représentation graphique des différences et similitudes observées dans l'alignement multiple des séquences des différentes espèces



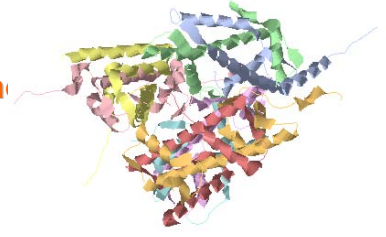
Arbre phylogénétique (H4)

H4, ou histone H4, est une protéine très importante pour la structure des chromosomes (enroulement de l'ADN autour des nucléosomes). Elle existe chez tous les organismes 'nucléés'.

Séquence en acides aminés de H4 humaine

```
MSGRGKGGKGLGKGGAKRHRKVLDRDNIQGITKPAIRRLARRGGVKRISGLIYEE
TRGVLVKFLENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGR
```

Structure 3D de H4 humaine



Alignement multiple des séquences de H4 des différentes espèces (extrait)

<input type="checkbox"/> Homme	MSGRGKGGKGLGKGGAKRHRKVLDRDNIQGITKPAIRRLARRGGVKRISGLIYEE
<input type="checkbox"/> Souris	MSGRGKGGKGLGKGGAKRHRKVLDRDNIQGITKPAIRRLARRGGVKRISGLIYEE
<input type="checkbox"/> Rat	MSGRGKGGKGLGKGGAKRHRKVLDRDNIQGITKPAIRRLARRGGVKRISGLIYEE
<input type="checkbox"/> Bœuf	MSGRGKGGKGLGKGGAKRHRKVLDRDNIQGITKPAIRRLARRGGVKRISGLIYEE
<input type="checkbox"/> Poulet	MSGRGKGGKGLGKGGAKRHRKVLDRDNIQGITKPAIRRLARRGGVKRISGLIYEE
<input type="checkbox"/> Grenouille	MSGRGKGGKGLGKGGAKRHRKVLDRDNIQGITKPAIRRLARRGGVKRISGLIYEE
<input type="checkbox"/> Ver	MSGRGKGGKGLGKGGAKRHRKVLDRDNIQGITKPAIRRLARRGGVKRISGLIYEE
<input type="checkbox"/> Mouche	MSGRGKGGKGLGKGGAKRHRKVLDRDNIQGITKPAIRRLARRGGVKRISGLIYEE
<input type="checkbox"/> Mais	MSGRGKGGKGLGKGGAKRHRKVLDRDNIQGITKPAIRRLARRGGVKRISGLIYEE
<input type="checkbox"/> Levure	MSGRGKGGKGLGKGGAKRHRKVLDRDNIQGITKPAIRRLARRGGVKRISGLIYEE

Cette protéine est particulièrement conservée au cours de l'évolution: chaque acide aminé a probablement une fonction essentielle à la Vie

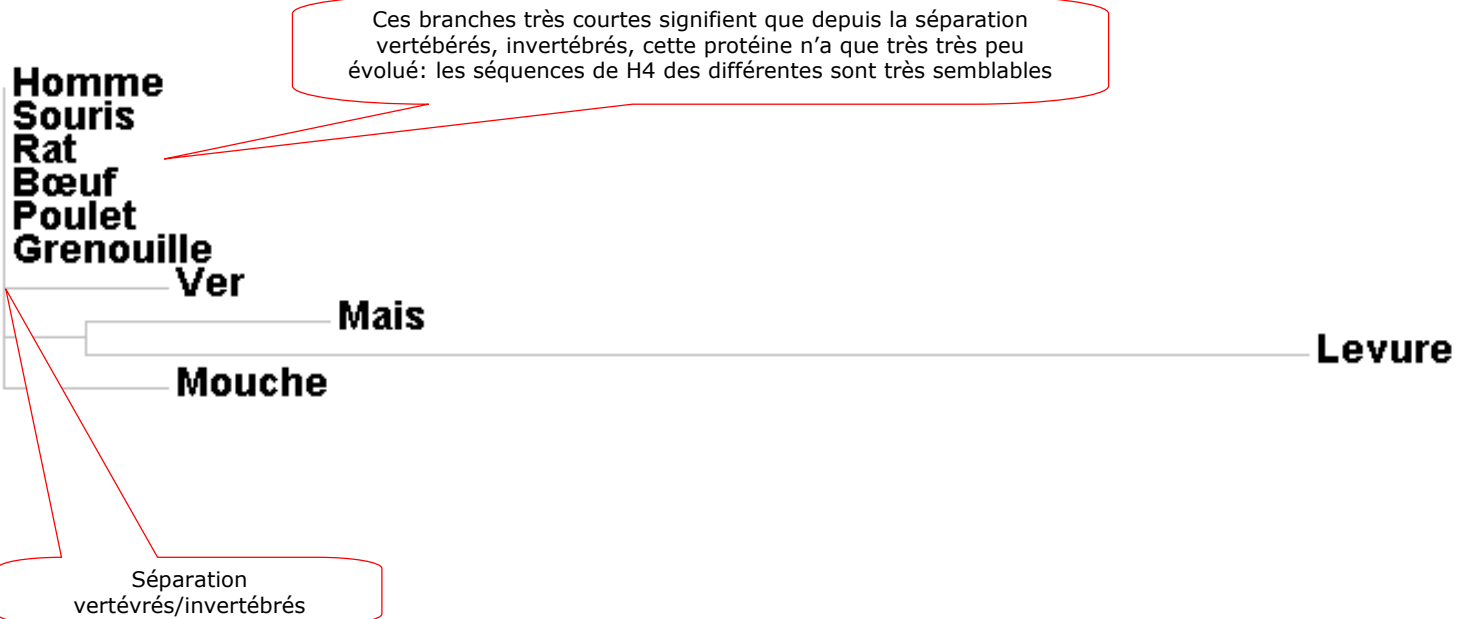
les colonnes colorées mettent en évidence les positions conservés entre les séquences des différentes espèces:
 ■ acides aminés identiques.
 ■ acides aminés avec les mêmes propriétés biochimiques.
 ■ acides aminés peu similaire du point de vue biochimique.
 ■ acides aminés non similaire du point de vue biochimique.

Calcul des différences entre les séquences



Arbre phylogénétique

Représentation graphique des différences et similitudes observées dans l'alignement multiple des séquences des différentes espèces



Arbre phylogénétique (ATPA)

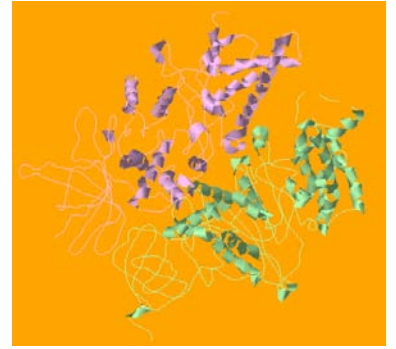
ATPA (ou ATP synthase alpha chain) est une protéine impliquée dans la production d'énergie.

Séquence en acides aminés de ATPA humaine

```

MLSVRVAADVRLPFRAGLVSRLNAGSSFIARNFHASNTHLQKTGTAEMSSILEERIL
GADTSVDLEETGRVLSIGDGIARVHGLRNRVQAEEMVEFSSGLKGMSLNLEPDNVGVVVF
NDKLIKEGDIVKRTGAIVDVVPVGEELLGRVVDALGNALDGGKPIGSKTRRRVGLKAPGII
PRISVREPMQTGIKAVDSLVPVIGRGQRELIIGDRQTGKTSIAIDTIINQKRPNDSDEKK
KLYCIYVAIGQKRSTVAQLVKRLTDADAMKVTIVVSATASDAAPLQYLAPYSGCSMGYF
RDNGKHALIYDDLKSKAVAYRQMSLLRRPPGREAYPGDVFYLSRLLERAARKMNDAPG
GGSLTALPVIETQAGDVSAYIPTNVISITDGGIFLETELPHYKIRPAINVGLSVSRVGS
AQTRAMKQVAGTMKLELAQYREVAAPAFQFSGSLLDAATQQLSRGVRLTELLKQGGYSPMA
IEEQVAVIYAGVRGYLDKLEPSKITKFNAPFLSHVVSQHALLGTIRADGKIIEQSDAKL
KEIVTNFLAGFEA
    
```

Structure 3D de ATPA de bactérie



Alignement multiple des séquences de ATPA des différentes espèces (extrait)

Espèce	Séquence
<input type="checkbox"/> Mais	QEGSFVKATGRIAQIPVSEAYLGRVINALAKPIDGRGEIVASESRLIESPAPGIIISRSVYELQGTGLIAIDSMPIIGRGQRELIIGDRQTGKTAVAT
<input type="checkbox"/> Riz	QEGSFVKATGRIAQIPVSEAYLGRVINALAKPIDGRGEIVASESRLIESPAPGIIISRSVYELQGTGLIAIDSMPIIGRGQRELIIGDRQTGKTAVAT
<input type="checkbox"/> Orang-outan	KEGDIVKRTGAIVDVVPVGEELLGRVVDALGNALDGGKPIGSKTRRRVGLKAPGIIPHISVREPMQTGIKAVDSLVPVIGRGQRELIIGDRQTGKTSIAI
<input checked="" type="checkbox"/> Homme	KEGDIVKRTGAIVDVVPVGEELLGRVVDALGNALDGGKPIGSKTRRRVGLKAPGIIPHISVREPMQTGIKAVDSLVPVIGRGQRELIIGDRQTGKTSIAI
<input type="checkbox"/> Bœuf	KEGDIVKRTGAIVDVVPVGEELLGRVVDALGNALDGGKPIGSKTRRRVGLKAPGIIPHISVREPMQTGIKAVDSLVPVIGRGQRELIIGDRQTGKTSIAI
<input type="checkbox"/> Rat	KEGDIVKRTGAIVDVVPVGEELLGRVVDALGNALDGGKPIGSKTRRRVGLKAPGIIPHISVREPMQTGIKAVDSLVPVIGRGQRELIIGDRQTGKTSIAI
<input type="checkbox"/> Souris	KEGDIVKRTGAIVDVVPVGEELLGRVVDALGNALDGGKPIGSKTRRRVGLKAPGIIPHISVREPMQTGIKAVDSLVPVIGRGQRELIIGDRQTGKTSIAI
<input type="checkbox"/> Grenouille	KEGDIVKRTGAIVDVVPVGEELLGRVVDALGNALDGGKPIGSKTRRRVGLKAPGIIPHISVREPMQTGIKAVDSLVPVIGRGQRELIIGDRQTGKTSIAI
<input type="checkbox"/> Mouche	KQGGDIKRTGAIVDVVPVGEELLGRVVDALGNALDGGKPIGSKTRRRVGLKAPGIIPHISVREPMQTGIKAVDSLVPVIGRGQRELIIGDRQTGKTALAI
<input type="checkbox"/> Ver	REGDVIKRTGAIVDVVPVGDGLLGRVVDALGNALDGGKPIANARRSRVEVKAPGIIPHSVREPMVTGVKAVDSLVPVIGRGQRELIIGDRQTGKTALAI
<input type="checkbox"/> Levure	KEGELVKRTGNIVDVVPVGGPLLGRVVDALGNALDGGKPIIDAGRSRAQVKAAPGIIPHSVHEEVQGTGLKAVDALVPIIGRGQRELIIGDRQTGKTAVAL
<input type="checkbox"/> Bactérie	AEGMKVKTGRILEVPVGRGLLGRVVDALGNALDGGKPLDHDGFSVAEIAAPGVIERQSVDFVQGTGYKAVDSLVPVIGRGQRELIIGDRQTGKTALAI

les colonnes colorées mettent en évidence les positions conservés entre les séquences des différentes espèces:

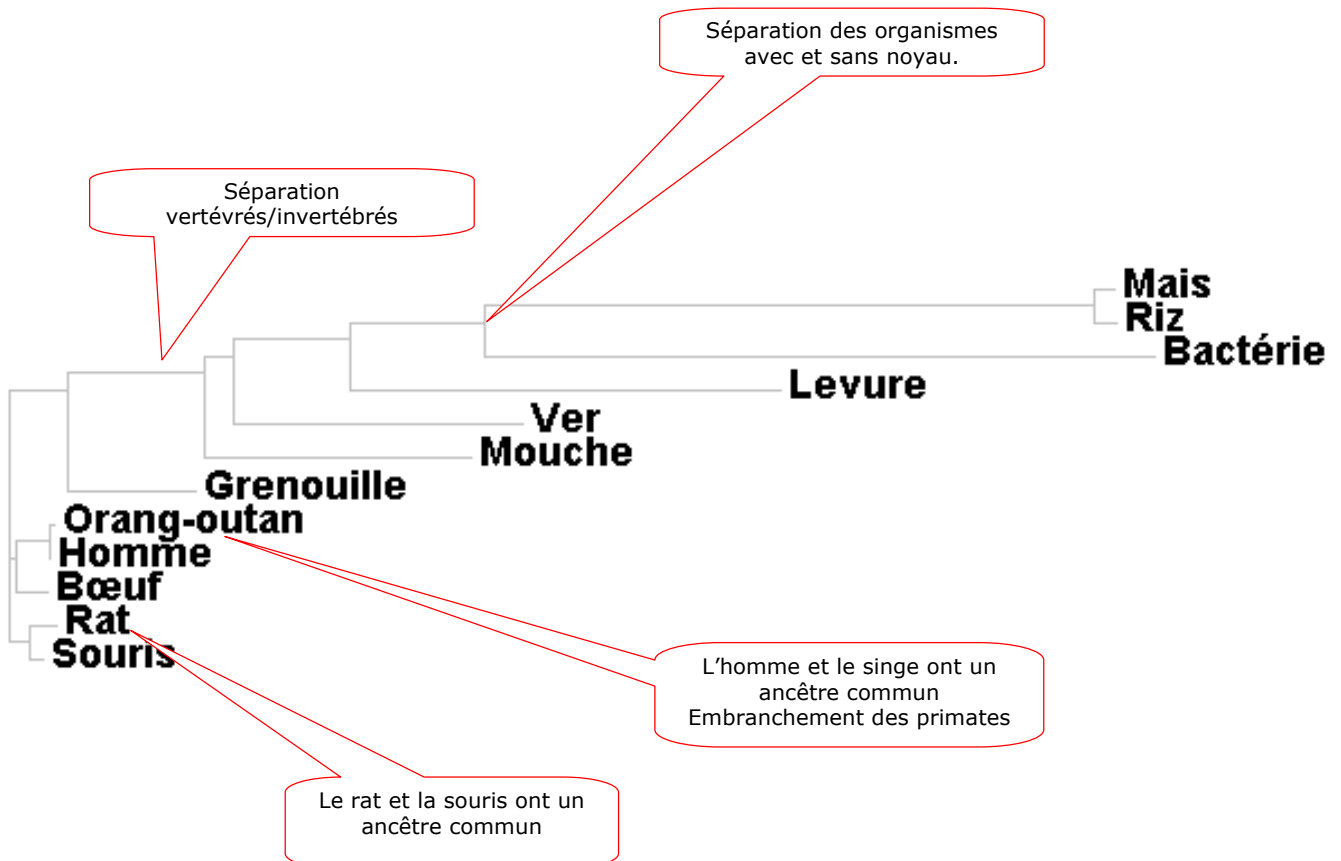
- acides aminés identiques.
- acides aminés avec les mêmes propriétés biochimiques.
- acides aminés peu similaire du point de vue biochimique.
- acides aminés non similaire du point de vue biochimique.

Calcul des différences entre les séquences



Arbre phylogénétique

Représentation graphique des différences et similitudes observées dans l'alignement multiple des séquences des différentes espèces



Arbre phylogénétique (hormone de croissance)

L'hormone de croissance (SOMA), comme son nom l'indique est une hormone indispensable à la croissance. Elle est synthétisée par l'hypophyse puis secrétée dans la circulation sanguine. Un manque de cette hormone conduit, entre autre, au nanisme.

Séquence en acides aminés de l'hormone de croissance humaine

```
MATGSRTSLLAFGLLCLPWLQEGSAFPTIPLSRLEFDNAMLRAHRLHQLAFDITYQEFEEA
YIPKEQKYSFLQNPQTSCLCFSESIPTPSNREETQQKSNLELLRISLLLIQSWLEPVQFLR
SVFANSLVYGASDSNVYDLLKDLLEEGIQTLMGRLEDGSPRTGQIFKQYTSKFDNTNSHND
ALLKNYGLLYCFRKMDKQVETFLRIVQCRSVEGSCGF
```

Structure 3D de l'hormone de croissance humaine



Alignement multiple des séquences des hormones de croissance des différentes espèces (extrait)

<input checked="" type="checkbox"/> Chimpanzé	-MAPGSRTSLLAFGLLCLPWLQEGSAFPTIPLSRLEFDNAMLRAHRLHQLAFDITYQEFEEA	AYIPKEQKYSFLQNPQTSCLCFSESIPTPSNREETQQK:
<input checked="" type="checkbox"/> Homme	-MATGSRTSLLAFGLLCLPWLQEGSAFPTIPLSRLEFDNAMLRAHRLHQLAFDITYQEFEEA	AYIPKEQKYSFLQNPQTSCLCFSESIPTPSNREETQQK:
<input checked="" type="checkbox"/> Rat	-MAADSQTPWLLTFSLLCLLWPQEGAFPPAMPSSLEFANAVLRAQHLHQLAADTYKEFER	AYIPEGORYS-IQNAQA AFCFSETI PAFTGKKEAQQR:
<input checked="" type="checkbox"/> Souris	-MATDSRTSWLLTVSLLCLLWPQEGAFPPAMPSSLEFANAVLRAQHLHQLAADTYKEFER	AYIPEGORYS-IQNAQA AFCFSETI PAFTGKKEAQQR:
<input checked="" type="checkbox"/> Lapin	-MAAGSWTAGLLAFALLCLPWPQEGAFPPAMPSSLEFANAVLRAQHLHQLAADTYKEFER	AYIPEGORYS-IQNAQA AFCFSETI PAFTGKDEAQQR:
<input checked="" type="checkbox"/> Bœuf	MMAAGPRTSLLAFALLCLPWTQVVGAFPPAMPSSLEFANAVLRAQHLHQLAADTYKEFER	TYIPEGORYS-IQNTQVAFCFSETI PAFTGKNEAQQR:
<input checked="" type="checkbox"/> Grenouille	-MATG-FCSSFGLLVVLLKKNVADVGAFPSVPLFSLEFNAVSRQYIHLMAADTYRDER	TYITDEQRHS-NKNSHVVCYSETI PFTDKDNTHQK:
<input checked="" type="checkbox"/> Poulet	-MAPGSWFSPLLIAVVTGLPQEAATFPAMPSSLEFANAVLRAQHLHLLAETTYKEFER	TYIPEGORYS-IQNTQVAFCFSETI PAFTGKDDAQQR:

les colonnes colorées mettent en évidence les positions conservés entre les séquences des différentes espèces:

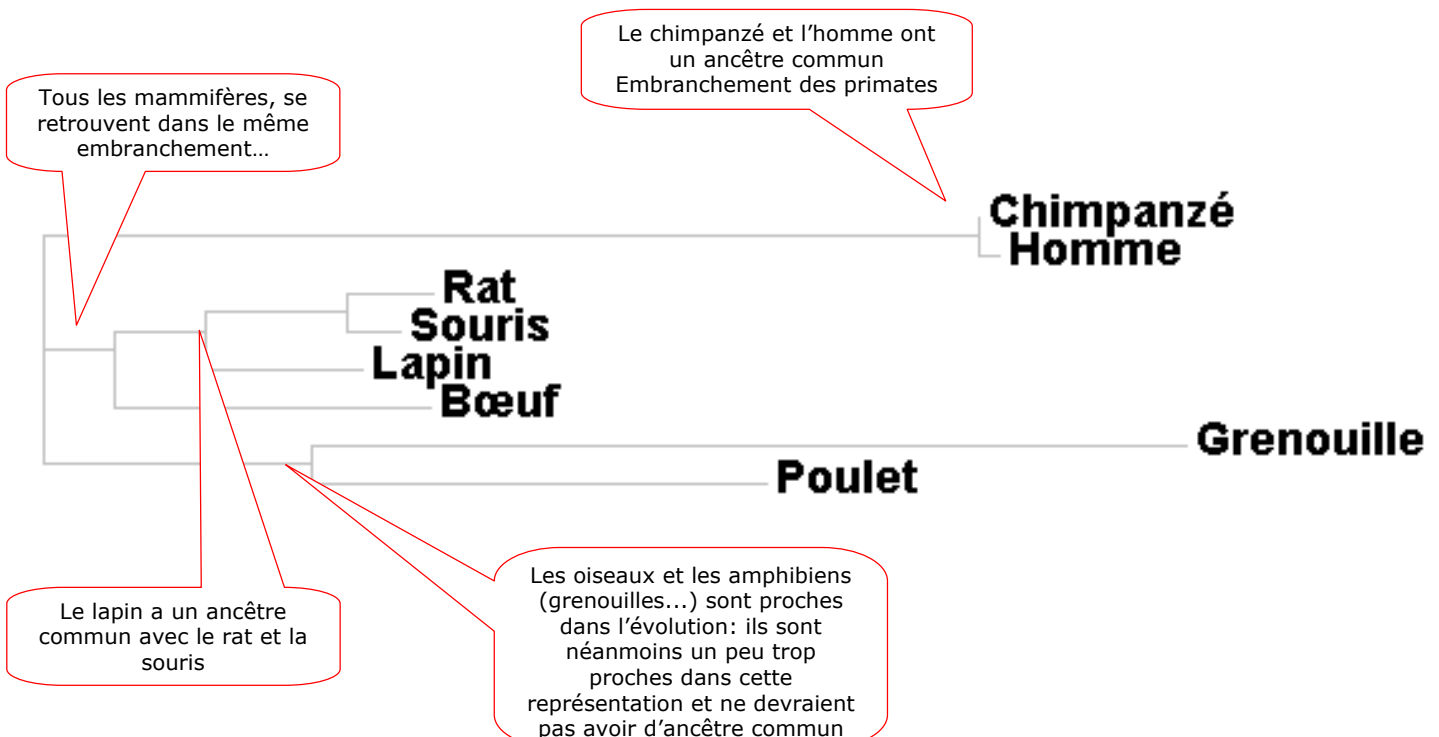
- acides aminés identiques.
- acides aminés avec les mêmes propriétés biochimiques.
- acides aminés peu similaire du point de vue biochimique.
- acides aminés non similaire du point de vue biochimique.

Calcul des différences entre les séquences



Arbre phylogénétique

Représentation graphique des différences et similitudes observées dans l'alignement multiple des séquences des différentes espèces



Arbre phylogénétique (cytochrome B)

Le cytochrome B est une protéine impliquée dans la production d'énergie. Elle est fabriquées dans les mitochondries, des organelles présentes en copie multiple dans les cellules. Conséquence: on retrouve suffisamment de matériel génétique dans les fossiles ou les restes d'animaux, ce qui permet de retrouver la séquence de cette protéine dans de nombreuses espèces (~1250), même éteintes (Dodo, Mammouth...)

Séquence en acides aminés du cytochrome B

MMCGAPSATQPATAETQHIADQVRSQLEEKENKFPVFKAVSFKSQVAVG
TNYFIKVHVG DEDFVHLRVFQSLPHENKPLTLSNYQTNAKAKHDELTYF

Structure 3D du cytochrome B bovin



Alignement multiple des séquences des cytochromes B des différentes especes (extrait)

☐ Mais	--MTIRNQRFSLLKQPIYSTLNQHLIDYPTPSNLSYWGFCCLAGICIVICIVVFLMHHTPHVLA	NSVEHIMRDVEGGWLLRYMANGASMFLIVVH
☐ Riz	--MTIRNQRFSLLKQPIYSTLNQHLIDYPTPSNLSYWGFCCLAGICIVICIVVFLMHHTPHVLA	NSVEHIMRDVEGGWLLRYMANGASMFFIVVY
☐ Levure	--MAFRKS-----NVYLSLVNSYIIDSPPSSINYNMGSLLGLCIVICIVVFLMHHTPHVLA	NSVEHIMRDVHNGYLLRYLANGASFFFMVMF
☐ Rat	--MTNIRKS-----HPLFKIINHSDLDLPPSPNISWVNFSSLLGVCIMVCIITLFLMHHTSDTMTA	SSVTHICRDVNYGWLIRYLANGASMFFICLF
☐ Souris	--MTNMRKT-----HPLFKIINHSDLDLPPSPNISWVNFSSLLGVCIMVCIITLFLMHHTSDTMTA	SSVTHICRDVNYGWLIRYLANGASMFFICLF
☐ Bœuf	--MTNIRKS-----HPLMKIVNNAFIDLPPSPNISWVNFSSLLGVCIMVCIITLFLMHHTSDTMTA	SSVTHICRDVNYGWLIRYLANGASMFFICLY
☐ Mammouth	--MTHIRKS-----HPLLLKLNKSFIDLPTSPNISTWVNFSSLLGACITCITLFLMHHTPDMTA	SSMSHICRDVNYGWLIRYLANGASIFFLCLY
☐ Eléphant	--MTHIRKS-----HPLLLKLNKSFIDLPTSPNISTWVNFSSLLGACITCITLFLMHHTPDMTA	SSMSHICRDVNYGWLIRYLANGASIFFLCLY
☐ Chimpanzé	--MTPMRKI-----NPLMKLNHSFIDLPTSPNISAWVNFSSLLGACITCITLFLMHHTPDASTA	SSIAHITRDVNYGWLIRYLANGASMFFICLF
☐ Homme	--MTPMRKI-----NPLMKLNHSFIDLPTSPNISAWVNFSSLLGACITCITLFLMHHTPDASTA	SSIAHITRDVNYGWLIRYLANGASMFFICLF
☐ Orang-outan	--MTPMRKT-----NPLMKLNHSFIDLPTSPNISAWVNFSSLLGACITCITLFLMHHTPDASTA	SSIAHITRDVNYGWLIRYLANGASMFFICLF
☐ Dodo	-----WVNFSSLLGICIMVCIITLFLMHHTADTTA	SSVAHTCRDNYGWLIRYLANGASFFFCICY
☐ Poulet	MAPNIRKS-----HPLMKLNHSFIDLPPSPNISAWVNFSSLLAVGIMTITLFLMHHTADTSLA	SSVAHTCRDNYGWLIRYLANGASFFFCICY
☐ Grenouille	MAPNIRKS-----HPLIKIINNSFIDLPTSPNISWVNFSSLLGVCIMVCIITLFLMHHTADTSMR	SSVAHICRDVNYGWLIRYLANGASFFFCICY
☐ Poisson	--MTSLRKT-----HPLVKIANDALVDLPTSPNISAWVNFSSLLGICITCITLFLMHHTADISTA	SSVNHICRDVNYGWLIRYLANGASFFFCICY
☐ Mouche	MNKPLRNS-----HPLFKIANNALVDLPPINISWVNFSSLLGICITCITLFLMHHTADINLA	YSVNHICRDVNYGWLIRYLANGASFFFCICY
☐ Ver	--MKINNS-----LLNFVNGMLVTLFSSKTLTLS	NFSSMLGMVILFQILHTELEFYITPDSIML

les colonnes colorées mettent en évidence les positions conservés entre les séquences des différentes espèces:

- ☐ acides aminés identiques.
- ☐ acides aminés avec les mêmes propriétés biochimiques.
- ☐ acides aminés peu similaire du point de vue biochimique.
- ☐ acides aminés non similaire du point de vue biochimique.

Calcul des différences entre les séquences



Arbre phylogénétique

Représentation graphique des différences et similitudes observées dans l'alignement multiple des séquences des différentes espèces

