

Emmanuel Douzery et Frédéric Delsuc

# Les horloges de l'évolution



Le temps laisse aussi son empreinte dans les molécules du vivant. Les biologistes exploitent ce phénomène, nommé horloge moléculaire, pour remonter le cours de l'évolution des organismes.

## L'ESSENTIEL

✓ Des changements surviennent dans l'ADN et les protéines au cours du temps. Le rythme de leur apparition, relativement constant, constitue l'horloge moléculaire.

✓ Cette horloge permet de dater la divergence des groupes phylogénétiques.

✓ En « assouplissant » l'horloge, on tient mieux compte des variations du rythme d'évolution de l'ADN d'une espèce à une autre.

© Shutterstock/stephan kerkhofs, aquatic creature, Fernando Cortes, Anyka, Hua Chee Kong, Pour la Science

**D**inosaures, ptérosaures, ammonites, bélemnites... Ces animaux ont disparu de la surface du globe il y a plus de 65 millions d'années. Comment le savons-nous ? Par une méthode éprouvée : la datation par isotopes radioactifs des terrains les plus récents dans lesquels leurs fossiles ont été découverts. On dispose ainsi d'un registre paléontologique qui s'étend jusqu'aux premières cellules, datées de 3,5 milliards d'années.

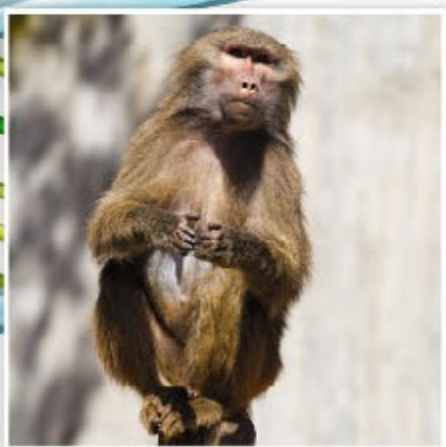
Le problème, c'est que les archives paléontologiques sont forcément incomplètes. Nous ne sommes jamais sûrs d'avoir mis au jour le plus vieux représentant d'un groupe donné. L'âge le plus reculé d'un groupe d'organismes n'est donc jamais certain. En outre, certains fossiles sont difficilement identifiables, et l'on ne peut les rattacher avec certitude aux lignées des organismes vivants actuels. Par ailleurs, beaucoup d'organismes ne se sont tout simplement pas fossilisés, notamment parce qu'ils n'étaient constitués que de tissus mous.

Aujourd'hui, pour décrire l'histoire de la vie, les spécialistes de l'évolution font

appel non seulement aux caractères morphologiques et anatomiques des organismes, mais aussi aux molécules, c'est-à-dire à l'ADN (du noyau cellulaire et des organites), aux ARN et aux protéines. L'étude de ces différents caractères permet de construire des arbres phylogénétiques.

Ces derniers sont constitués de branches, qui symbolisent l'évolution des lignées d'espèces, et de nœuds, qui désignent leurs ancêtres communs. Il en ressort deux types d'informations : les relations de parenté des organismes (deux lignées liées par un nœud sont étroitement apparentées), et leur « quantité d'évolution », représentée par la longueur des branches de l'arbre.

Sur ces arbres, nous pouvons remonter le temps et déterminer quand les groupes d'organismes ont divergé, en profitant d'un phénomène découvert dans les années 1960 : l'horloge moléculaire. Voyons comment elle fonctionne et comment on s'en sert.



Les nucléotides et les acides aminés, les constituants respectifs de l'ADN et des protéines, changent naturellement dans les cellules. Or, en 1962, le biologiste austro-américain Emile Zuckerkandl et le biochimiste américain Linus Pauling, de l'Institut de technologie de Californie, proposèrent que ces changements se produisent au cours du temps à un rythme constant quelle que soit l'espèce. Ils avaient constaté que la séquence (la suite des acides aminés) d'une protéine, la globine, présente chez l'homme 18 différences comparée à celle du cheval. Ils connaissaient par ailleurs, d'après les fossiles, à quelle époque la branche menant aux primates s'était séparée de celle des périssodactyles, le taxon du cheval. Ils conclurent à l'existence d'une horloge moléculaire dont la valeur était égale à un certain nombre de changements d'acides aminés par million d'années ; ils définirent ainsi un taux d'évolution. Ce principe pouvait être généralisé à d'autres espèces (voir l'encadré page 104).

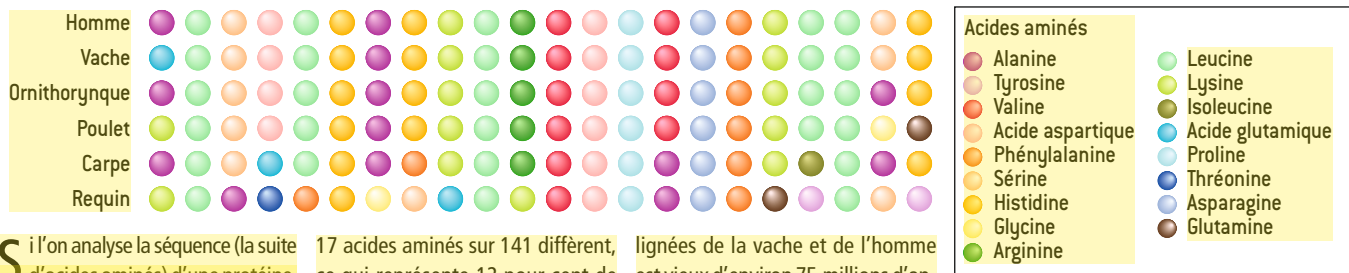
## Un chronomètre dans l'ADN

Admettons que les substitutions de nucléotides dans l'ADN, et d'acides aminés dans les protéines, apparaissent à un rythme constant, quelle que soit l'espèce : c'est l'hypothèse de l'horloge moléculaire globale, ou stricte, proposée par Zuckerkandl et Pauling. Il devient alors possible de déduire une date de séparation des lignées des deux espèces du nombre de changements moléculaires les distinguant – ou divergence moléculaire. Cela revient à estimer l'âge de leur dernier ancêtre commun. Dans certains cas, on retrouve la tranche d'âges établie à l'aide des fossiles. Dans d'autres, lorsque les archives fossiles sont restreintes ou absentes, on comble, en utilisant cette méthode, un vide dans l'histoire de l'évolution.

Par exemple, les plus proches parents des oiseaux, dans la nature actuelle, sont les crocodiliens. Or une protéine nommée alpha-globine diffère de 28 pour cent entre le poulet et l'alligator. Dans

**1. L'HORLOGE MOLÉCULAIRE** correspond au nombre de changements survenant dans une molécule au cours d'un temps donné. Associée à la description des relations de parenté entre organismes (de haut en bas, un ver plat, une mouche, une grenouille et un singe), elle permet de dater le moment où ces lignées ont divergé.

### SÉQUENCE PARTIELLE DE L'ALPHA-GLOBINE



**S**i l'on analyse la séquence (la suite d'acides aminés) d'une protéine, quelle qu'elle soit, on constate qu'elle diffère d'une espèce à l'autre. Ainsi en est-il des protéines de la famille des globines, qui assurent le transport de l'oxygène chez les vertébrés. L'alpha-globine constitue l'une des quatre chaînes protéiques possibles (avec les chaînes bêta, gamma et delta) de la molécule d'hémoglobine, présente dans les globules rouges du sang. La comparaison de sa séquence chez l'homme et la vache montre que

17 acides aminés sur 141 différent, ce qui représente 12 pour cent de divergence (*ci-dessus*). Comparée à l'alpha-globine de l'ornithorynque, du poulet, de la carpe et du requin, celle de l'homme diverge de 27, 30, 50 et 56 pour cent, respectivement.

L'horloge moléculaire repose sur le nombre de changements d'acides aminés qui se produisent au cours du temps. Pour étalonner cette horloge, il faut faire appel au registre fossile des vertébrés. Par exemple, le dernier ancêtre commun connu des

lignées de la vache et de l'homme est vieux d'environ 75 millions d'années ; celui des lignées du requin et de l'homme a environ 430 millions d'années. En plaçant les différents ancêtres communs sur une échelle de temps, on peut dater les nœuds d'un arbre phylogénétique décrivant les relations de parenté entre les six espèces (*ci-contre, a*).

En représentant l'âge de l'ancêtre commun connu (*en abscisses*) en fonction du pourcentage de divergence de l'alpha-globine de l'homme

vis-à-vis des cinq autres espèces (*ci-contre, b, en ordonnées*), on constate que la divergence augmente avec le temps. La valeur de l'horloge, quant à elle, reste, en moyenne, constante durant 500 millions d'années (*droite verte*) : elle est ici de 0,125 pour cent par million d'années. Comme il s'agit d'une moyenne, il existe des écarts avec les âges de certains ancêtres communs estimés d'après les fossiles (*flèches rouges*).

l'hypothèse de l'horloge globale, l'âge de la séparation des deux lignées, donc l'âge de leur dernier ancêtre commun, peut être calculé, sachant que 0,125 pour cent de changements surviennent par million d'années. Il serait ici de  $28/0,125$ , soit 224 millions d'années. Les archives fossiles, quant à elles, font remonter les archosaures, les ancêtres des crocodiliens et des oiseaux, à environ 250 millions d'années.

## Les soubresauts de l'horloge

En pratique, l'utilisation de l'horloge moléculaire se révèle compliquée. Il est d'abord difficile, contrairement aux apparences, de déterminer exactement le degré de divergence d'une séquence moléculaire entre deux espèces. En effet, dans la molécule d'ADN, un nucléotide remplace d'autant plus facilement un autre nucléotide qu'il appartient à la même famille chimique (les bases puriques ou les bases pyrimidiques). De même, dans une protéine, un acide aminé hydrophobe est remplacé plus souvent par un autre de même propriété chimique – une glycine par une valine, par exemple – que par un acide aminé hydrophile, telle une sérine. Il

faut donc corriger les estimations de degré de divergence à l'aide de modèles de probabilité appropriés.

Deuxièmement, tous les changements de séquences ne sont pas détectés puisqu'un même site, dans une molécule, peut changer de multiples fois au cours de l'évolution, seul l'état final de la séquence étant observé. On risque donc de sous-estimer le taux de divergence moléculaire, et donc de faire passer une séparation



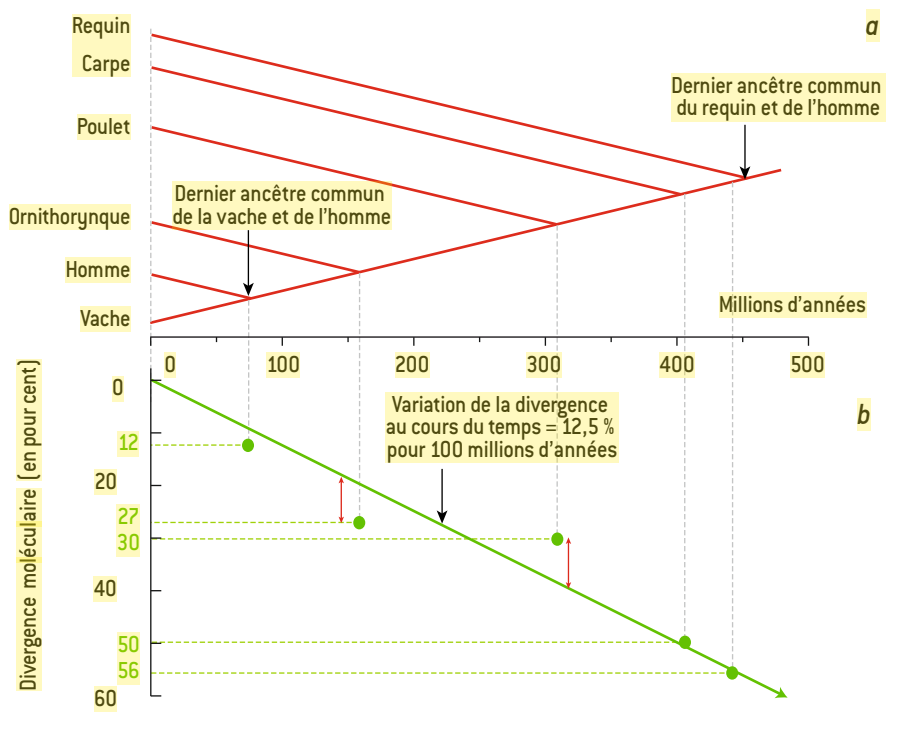
**2. MEMBRE DU GROUPE DES TUNICIERS**, cette ascidie (*Styela truncata*) fait partie, comme les vertébrés, de l'embranchement des chordés. Pourtant, le génome des tuniciers change plus rapidement que celui des vertébrés, ce qui disqualifie la notion d'horloge moléculaire à taux fixe.

de deux lignées pour plus récente qu'elle ne l'a été en réalité.

Plus problématique, l'hypothèse de l'horloge moléculaire globale ne tient pas si on la confronte à la réalité. Ainsi, lorsque nous construisons les arbres phylogénétiques des rongeurs et des primates, nous constatons que, pendant un même laps de temps, les génomes du rat et de la souris ont plus changé que ceux de l'homme et du chimpanzé. De même, à une plus grande échelle taxonomique, l'ADN mitochondrial et nucléaire des tuniciers (violets, ciones, ascidies) change plus rapidement que celui de leurs plus proches parents, les vertébrés.

Ces disparités du rythme d'évolution des génomes ont des causes multiples. Citons, par exemple, la plus ou moins grande efficacité de la réparation de l'ADN (des espèces qui réparent mal leurs mutations évoluent plus vite), et la fréquence de la reproduction (des espèces qui se reproduisent plus fréquemment que d'autres sont sujettes à plus de mutations).

De surcroît, au sein d'une même espèce, différents gènes ou protéines peuvent évoluer à des vitesses différentes. En règle générale, plus une molécule est importante pour le fonctionnement de l'or-



## LES AUTEURS



Emmanuel DOUZERY, enseignant-chercheur à l'Université de Montpellier 2, dirige l'Équipe de phylogénie moléculaire, où Frédéric DELSUC est chargé de recherche CNRS, à l'Institut des sciences de l'évolution de Montpellier (UMR 5554 CNRS).

## ✓ SUR LE WEB

<http://www.timetree.org/>

ganisme, moins elle évolue vite. Par exemple, les histones, protéines liées à l'ADN, évoluent très lentement, tandis que les globines, protéines impliquées dans le transport de l'oxygène, et les fibrinopeptides, qui interviennent dans la coagulation sanguine, changent plus rapidement. Il n'existe donc pas une, mais des horloges moléculaires, selon la molécule étudiée... C'est un avantage : lorsque le taux d'évolution de la protéine est important, l'horloge représente un sablier utilisable pour étudier les courtes périodes ; lorsque le taux est faible, elle peut être l'équivalent d'horloges isotopiques pour les longues durées.

## Assouplir l'horloge

Toutefois, dès que l'on veut comparer une même molécule dans différentes lignées, on est confronté à la disparité des vitesses d'évolution moléculaire selon les espèces. Depuis quelques années, les biologistes ont mis au point des méthodes qui prennent en compte ces variations : celle des horloges locales, et celle des horloges assouplies, ou relâchées.

La première méthode suppose que les changements moléculaires se produisent à taux constant sur certaines parties de l'arbre

phylogénétique, celles reliant les groupes les plus proches – donc sur une période de temps donnée, de quelques millions d'années par exemple –, mais que leur rythme varie dans d'autres régions de l'arbre ou à de plus grandes échelles de temps. Il est cependant difficile de délimiter précisément les régions de l'arbre évoluant au même rythme, ainsi que le nombre de régions distinctes.

Les spécialistes préfèrent utiliser aujourd'hui une deuxième méthode, proposée en 1997 par Michael Sanderson, de l'Université de Californie à Davis, et développée depuis sous diverses formes. Elle présuppose que les vitesses d'évolution ne varient pas trop brusquement le long des branches d'un arbre phylogénétique. En effet, le taux d'évolution d'une branche de l'arbre phylogénétique est proche de celui de la branche dont elle descend.

Autrement dit, le taux de la branche parentale et celui des branches filles sont, en moyenne, identiques, comme dans le cas de l'horloge globale, mais avec des variations. Dès lors, le taux moyen peut prendre *a priori* une infinité de valeurs possibles. Heureusement, des méthodes mathématiques permettent de choisir des valeurs telles que la variation entre deux

## Trois critères de fiabilité

**Dater l'émergence des espèces à l'aide des horloges moléculaires requiert au moins trois conditions.**

✓ **Premièrement, il faut incorporer à l'analyse suffisamment d'espèces. On détecte ainsi davantage de substitutions de nucléotides ou d'acides aminés, notamment les changements qui se sont produits à plusieurs reprises sur le même site moléculaire.**

✓ **Deuxièmement, comme ce sont les fossiles qui nous fournissent les références de temps nécessaires, il est prudent d'utiliser plusieurs sources paléontologiques indépendantes de façon à disposer de plusieurs étalonnages des intervalles d'évolution.**

✓ **Troisièmement, plusieurs gènes ou protéines indépendants doivent être inclus dans les analyses. Avec un large échantillonnage, les datations moléculaires dépendent moins de particularités évolutives d'un gène ou d'une protéine.**