

# les apports récents de la paléoprotéomique à la paléoanthropologie : le cas des paranthropes

publié le 11.03.25 | par [damien aubert pennequin](#)

la paléoprotéomique, une discipline émergente qui analyse les protéines anciennes, offre de nouvelles perspectives fascinantes en paléoanthropologie, l'étude des hominidés éteints. en 2023, une équipe internationale a réussi à séquencer des peptides vieux de 2 ma, extraits de dents de *paranthropus robustus*. ces séquences révèlent des informations sur la diversité génétique et le sexe des individus étudiés, ainsi que sur la position phylogénétique de cette espèce dans la lignée humaine.

la compréhension de l'évolution humaine est un domaine riche qui combine diverses disciplines scientifiques. les techniques moléculaires viennent aujourd'hui compléter les méthodes plus traditionnelles d'analyse morphologique des fossiles. fin 2013, le séquençage presque complet du génome de l'homme de néandertal [1] a mis en lumière de manière spectaculaire [la toute jeune discipline de la paléogénomique](#) en montrant qu'elle pouvait apporter de précieuses informations sur notre histoire ancienne, par exemple en apportant des preuves d'hybridations entre humains archaïques et modernes, ou encore en permettant la découverte d'une toute nouvelle espèce, l'homme de denisova, à partir de seulement quelques dents et d'un petit bout de phalange [2]. cet exploit technique a ainsi valu à svante pääbo de recevoir en 2022 le prix nobel de physiologie ou médecine « pour ses découvertes sur les génomes d'espèces humaines éteintes et l'évolution humaine » [3].

contrairement à l'adn, qui se dégrade rapidement (figure 1), les protéines se révèlent plus résistantes et sont ainsi utilisées en paléoprotéomique, une discipline qui a pour objectif de faire parler les protéines anciennes, en particulier quand l'adn ancien a disparu ou est devenu complètement inexploitable. c'est par exemple par l'analyse des protéines anciennes que la mandibule de xiahe - trouvée dans la grotte karstique de baishiya dans le district de xiahe, dans la province de gansu, en chine - a pu être attribuée à l'homme de denisova, alors qu'aucune trace d'adn ancien n'avait été préservée [4]. la mandibule de xiahe a été datée d'environ 160 000 ans, mais le champ d'étude de la paléoprotéomique s'est dernièrement étendu à des restes fossiles bien plus anciens, tels que ceux du genre *paranthropus* [5a].

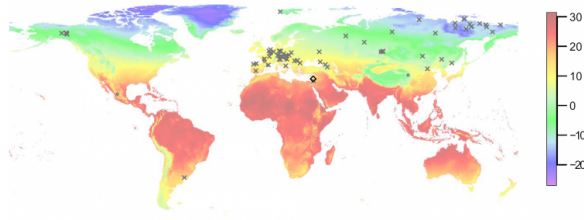


figure 1 - carte des sites où de l'adn de mammifères plus ancien que 30 ka a été retrouvé les croix marquent les sites où l'adn a été extrait de squelettes, les ronds marquent les sites où l'adn a été extrait de sédiments. le losange noir marque la grotte de sefunim. les fragments d'adn les plus anciens jamais séquencés viennent du groenland et datent de 2 ma. en dehors du permafrost il n'est pas possible de récupérer de l'adn plus ancien que 400 ka environ. les couleurs de la carte montrent les températures moyennes (°c) telles que mesurées entre les années 1970 et 2000.

auteur(s)/autrice(s) : slon et coll., 2022 licence : [cc-by](#)  
 source : [scientific reports](#)

les paranthropes forment un groupe d'hominidés éteints ayant vécu en afrique il y a 1 à 2,8 millions d'années. trois espèces sont généralement distinguées : *p. robustus*, *p. boisei* et *p. aethiopicus* (mais il est possible que ces deux dernières soient en réalité la même espèce). les paranthropes sont connus pour leur morphologie robuste, avec notamment une crête sagittale proéminente, des mâchoires puissantes et des dents très larges, avec un émail très épais, adaptées à un régime alimentaire herbivore abrasif (figure 2). les relations phylogénétiques entre les paranthropes et les autres hominidés, tels que les australopithèques et les premiers membres du genre *homo*, sont encore sujettes à débat. le genre *paranthropus*, avec ses caractéristiques morphologiques uniques et sa position controversée au sein de la lignée humaine, offre un terrain idéal pour illustrer l'apport des nouvelles technologies à la recherche paléanthropologique. c'est dans ce contexte que s'inscrivent les travaux menés par palesa p. madupe, claire koenig, ioannis patramanis et leurs collègues sur une population ancienne de *paranthropus robustus* [5b].

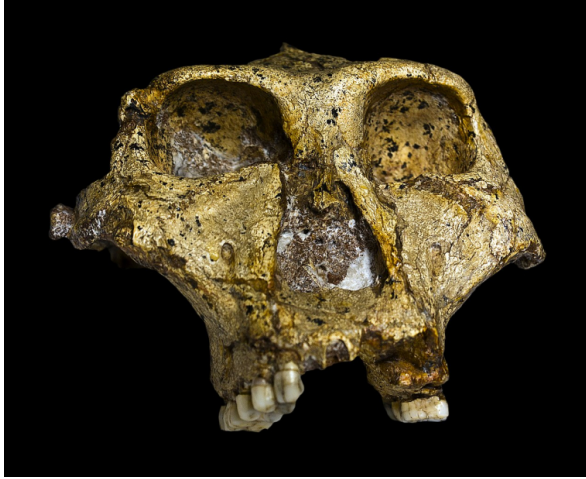


figure 2 - crâne original de *paranthropus robustus* le crâne original complet (sans la mandibule) de 1,8 million d'années de *paranthropus robustus* (sk-48 provenant de swartkrans (26°00's 27°45'e), gauteng), découvert en Afrique du Sud. collection du *transvaal museum*, northern flagship institute, Pretoria, Afrique du Sud.

auteur(s)/autrice(s) : José Braga et Didier Descouens  
licence : [cc-by-sa](https://creativecommons.org/licenses/by-sa/4.0/) source : [wikimedia](https://commons.wikimedia.org/wiki/File:Paranthropus_robustus_SK48.jpg)

## 1. Matériels et méthodes

L'objectif principal de l'étude de Madupe et collaborateurs est de clarifier les relations évolutives entre *paranthropus robustus* et les autres hominidés éteints, en complétant les analyses morphologiques traditionnelles par des données moléculaires. Pour ce faire, les chercheurs ont analysé quatre spécimens dentaires (figure 3) attribués à *p. robustus* trouvés sur le site de Swartkrans et datant d'environ 2 millions d'années (étage du Gelasien, pléistocène inférieur).

**figure 3 - les quatre dents de  
paranthropus robustus  
analysées par madupe et coll.,  
2023**

**les spécimens ont été trouvés  
dans la grotte de swartkrans en  
afrique du sud à 40 km au  
nord-ouest de johannesbourg.  
les dépôts sédimentaires  
contenant ces fossiles sont  
datés d'environ 1,8 à 2,2 ma.  
ces spécimens étaient  
conservés au musée national  
ditsong d'histoire naturelle. de  
gauche à droite : sk 835, une 3<sup>e</sup>  
molaire supérieure gauche ;  
sk 830 une 4<sup>e</sup> prémolaire  
inférieure gauche ; sk 850, une  
3<sup>e</sup> prémolaire inférieure  
droite ; sk 14132, une 3<sup>e</sup>  
molaire supérieure droite.  
notez que les deux prémolaires  
humaines sont numérotées 3 et  
4 parce que les prémolaires 1  
et 2 ont été perdues au cours  
de l'évolution.**



**auteur(s)/autrice(s) : madupe  
et coll., 2023 licence : [cc-by-  
nc-nd](#) source : [biorxiv](#)**

**l'émail dentaire a été réduit en poudre et dissout avec de l'acide trifluoroacétique pour en extraire les peptides anciens. cette opération a été réalisée à l'université de copenhague dans un laboratoire spécialement équipé pour réduire les risques de contamination par des peptides modernes. après nettoyage, les peptides ont été séparés par chromatographie en phase liquide, puis analysés par spectrométrie de masse en tandem (lc-ms/ms). cette technique consiste dans un premier temps à ioniser les peptides, puis à mesurer leurs rapports masse/charge. dans un deuxième temps les peptides sont fragmentés par des collisions avec un gaz inerte et les rapports masse/charge de ces fragments sont à leur tour mesurés (d'où le terme « en tandem »). le traitement des données ainsi récoltées permet de déterminer la séquence du peptide analysé.**

**ces séquences peptidiques ont ensuite été minutieusement comparées et alignées avec les protéines connues de l'émail dentaire chez différentes espèces de singes plus ou moins éloignées de l'homme moderne. des arbres phylogénétiques ont été produits pour chaque protéine individuellement et pour une concaténation de toutes les protéines (c'est-à-dire toutes les séquences alignées mises bout à bout). les chercheurs n'ont pas utilisé le critère traditionnel du maximum de parcimonie pour construire ces arbres mais deux méthodes probabilistes plus sophistiquées : le maximum de vraisemblance et l'inférence bayésienne (voir l'encadré ci-dessous). l'étude inclut également une analyse morphométrique des dents de swartkrans ainsi que de dents provenant d'autres sites.**

## les critères d'optimalité d'un arbre phylogénétique

il existe différents critères permettant d'évaluer la justesse de plusieurs arbres phylogénétiques en concurrence et ainsi de choisir le meilleur. dans les cas les plus complexes, il est nécessaire de recourir aux critères les plus sophistiqués.

### 1. maximum de parcimonie (mp)

le principe de parcimonie énonce que le meilleur arbre est celui qui explique les données (les séquences observées) avec le plus petit nombre total de changements évolutifs (les mutations). le scénario le plus parcimonieux est privilégié, ce qui revient à chercher l'arbre le plus court si les longueurs des branches sont proportionnelles au nombre de transformations qu'elles accueillent.

avantages : assez rapide, facile à comprendre.

limites : non fiable. il est malheureusement démontré théoriquement et empiriquement que ce critère n'est pas statistiquement cohérent [6][7][8]. cela signifie que les méthodes utilisant ce critère convergent vers un arbre incorrect même si la quantité de données disponibles tend vers l'infini. cela est dû à différents biais statistiques, liés notamment au fait qu'une même position peut muter plusieurs fois au cours de l'évolution. la quantité d'homoplasies est ainsi systématiquement sous-estimée.

### 2. minimum d'évolution (me)

le principe d'évolution minimale ressemble beaucoup au critère précédent : le meilleur arbre est le plus court [9]. mais les longueurs des branches ne sont pas ici estimées par le nombre minimum de mutations accommodant la topologie de l'arbre. les longueurs des branches sont calculées à partir des distances génétiques entre toutes les paires de séquences, et ces distances génétiques sont elles-mêmes le résultat d'une estimation statistique du nombre de mutations qui se sont *réellement* produites le long du chemin reliant deux séquences, et non simplement le nombre de mutations observées. ces estimations des distances génétiques reposent sur des modèles stochastiques d'évolution des séquences. le *neighbor-joining* est l'algorithme heuristique le plus connu implémentant ce principe d'évolution minimale, ou plus exactement le critère bme (*balanced minimum evolution*). le critère bme correspond à l'ajout d'une pondération des distances génétiques lors des calculs des longueurs des branches pour une topologie particulière, ce qui augmente les performances [10].

avantages : efficace et très rapide pour traiter de grandes quantités de données.

limites : bien que fiable (statistiquement cohérent) et robuste (résistant à d'importantes perturbations des données), ce critère n'est pas aussi puissant que des méthodes statistiques plus sophistiquées.

### 3. maximum de vraisemblance (ml)

la méthode du maximum de vraisemblance cherche l'arbre qui maximise la probabilité d'observer les données (les séquences) avec un certain modèle d'évolution [11][12][13]. l'adéquation du modèle aux données est elle-même évaluée par cette méthode d'inférence phylogénétique. le maximum de vraisemblance est un critère bien connu des statisticiens, mais son interprétation est peu intuitive. il s'agit en effet d'un calcul de la *probabilité des données*, et non de la probabilité de l'arbre. cela ne garantit pas que l'arbre choisi est le plus probable. en particulier, la somme des vraisemblances de tous les arbres possibles n'est pas égale à 1. autrement dit, cette méthode repose sur le paradigme fréquentiste, et cela pose les mêmes problèmes d'interprétation qu'une valeur  $p$ , ou qu'un intervalle de confiance par exemple. calculer l'intervalle de confiance à 95 % d'une variable  $x$  ne signifie pas que  $x$  a 95 % de chances d'être dans cet intervalle. cela signifie que, si l'on répétait l'expérience de nombreuses fois et calculait un nouvel intervalle de confiance à chaque fois, 95 % de ces intervalles contiendraient

la valeur réelle de  $x$ .

**avantages** : calculs probabilistes explicites et réalistes prenant en compte toutes les informations disponibles.

**limites** : plus complexe et coûteux en temps de calcul, paradigme fréquentiste rendant l'interprétation difficile.

#### 4. maximum à posteriori (map) / inférence bayésienne

dans le paradigme bayésien on cherche l'arbre ayant la probabilité la plus élevée. cette probabilité est dite à posteriori, car elle est calculée en tenant compte à la fois des données observées et d'une distribution à priori des probabilités. la somme des probabilités de tous les arbres possibles est bien égale à 1. de plus, cette méthode intègre l'incertitude sur les paramètres du modèle en les marginalisant, c'est-à-dire en considérant toutes les valeurs possibles de ces paramètres pondérées par leur probabilité : taux de mutation, longueurs de branches, etc. c'est un avantage considérable par rapport au maximum de vraisemblance puisqu'avec cette dernière on ne calcule pas la probabilité des données étant donné la topologie de l'arbre, mais la probabilité des données étant donné la topologie et les valeurs optimisées des différents paramètres. c'est ce qu'on appelle une probabilité conjointe. or la détermination de ces paramètres n'est pas l'objectif recherché. le but poursuivi est d'utiliser ces paramètres pour déterminer la topologie de l'arbre et c'est ce tour de force que parvient à effectuer l'inférence bayésienne : extraire la probabilité d'une topologie seule pour toutes les distributions estimées de tous les paramètres. mais cette prouesse n'est possible que grâce à de très lourdes simulations, il est absolument impossible de calculer ces probabilités de manière analytique [14].

**avantages** : évaluation explicite de la probabilité marginale d'un arbre rendant l'interprétation simple, incorpore toutes les données disponibles dans un modèle réaliste.

**limites** : exige des temps de calcul considérables reposant sur des simulations.

## 2. résultats et discussion

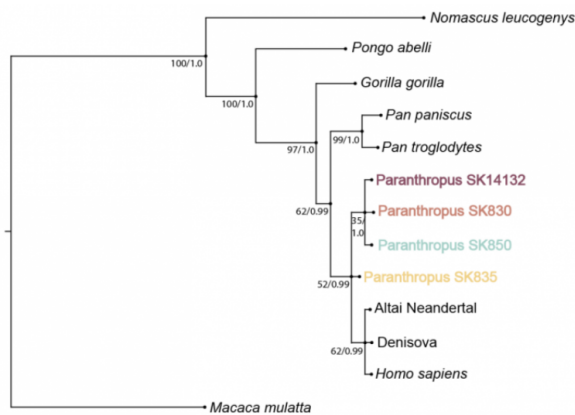
l'extraction des protéines de paranthropes et la détermination de leurs séquences a été un succès. le transport des restes au cours de crues subites et le dépôt des sédiments au sein d'une grotte dans une zone au climat aride a certainement contribué à leur bonne préservation en limitant les oscillations thermiques. les protéines ainsi récupérées sont principalement des protéines structurales de la matrice extracellulaire de l'émail dentaire telles que l'amélogénine, l'améloblastine, l'énaméline, le collagène, etc. l'analyse détaillée des séquences montre des différences génétiques entre les individus et même de l'hétérozygotie (deux allèles différents chez un même individu).

il a également été possible d'identifier le sexe des quatre spécimens. en effet, l'amélogénine possède deux isoformes, l'une produite par le gène *amelx* porté par le chromosome x, l'autre produite par le gène *amely* porté par le chromosome y. l'amélogénine-x et l'amélogénine-y n'ont pas tout à fait la même séquence et peuvent être distinguées. comparez par exemple ce peptide dans l'amélogénine-x « yqsirppyps » avec celui correspondant dans l'amélogénine-y « yqsmirppys ». l'identification de ces peptides spécifiques de l'amélogénine-y chez deux spécimens a permis de les identifier comme des mâles (sk 850 et sk 835) tandis que les deux autres sont très probablement des femelles (sk 830 et sk 14132). l'individu sk 835 avait précédemment été identifié comme une femelle sur la base de sa petite taille, ce qui met en lumière les limites des méthodes de sexage fondées sur le dimorphisme sexuel. il faut toutefois noter que le

**gène *amely* est parfois absent (délétion complète) chez certains mâles comme cela a été rapporté pour *homo sapiens* et pour *homo neanderthalensis*. cette technique moléculaire pourrait donc faussement identifier des mâles comme étant des femelles dans de rares cas.**

**figure 4 - arbre phylogénétique reconstruit par madupe et coll., 2023**

les nombres aux nœuds représentent à gauche le *bootstrap* de 0 à 100 des répliquats de l'analyse par maximum de vraisemblance et à droite la probabilité a posteriori de 0 à 1 de l'analyse par inférence bayésienne. le *bootstrap* est une mesure permettant d'évaluer la stabilité d'un clade lorsque l'analyse est effectuée à plusieurs reprises (généralement au moins 1000 fois). pour chaque répliquat on rééchantillonne aléatoirement les données, c'est-à-dire qu'on tire au sort (avec remise) les positions des résidus dans l'alignement des séquences qui sont conservées ou éliminées. cela revient à perturber les données en attribuant un poids aléatoire à ces différentes positions, puisque certaines peuvent ne pas être choisies alors que d'autres peuvent être choisies une ou plusieurs fois. on observe alors le pourcentage de répliquats dans lesquels le meilleur arbre contient le clade d'intérêt. un score de *bootstrap* faible (inférieur à 80 %) est le signe d'une grande sensibilité des données aux perturbations. cela ne signifie pas que ces clades sont nécessairement faux, ni même improbables, mais cela rend leur position moins fiable, on ne peut pas leur accorder la même confiance qu'à ceux qui sont plus stables.



auteur(s)/autrice(s) : madupe et coll., 2023 licence : [cc-by-nc-nd](https://creativecommons.org/licenses/by-nc-nd/4.0/) source : [biorxiv](https://www.biorxiv.org/)

la reconstruction phylogénétique (figure 4) montre que tous les spécimens de *p. robustus* sont plus proches du genre *homo* que des autres grands singes, confirmant ainsi l'appartenance des paranthropes à la lignée humaine. ce résultat est bien sûr cohérent avec les meilleures phylogénies de la lignée humaine disponibles actuellement, comme celle présentée dans la figure 5 [15][16].

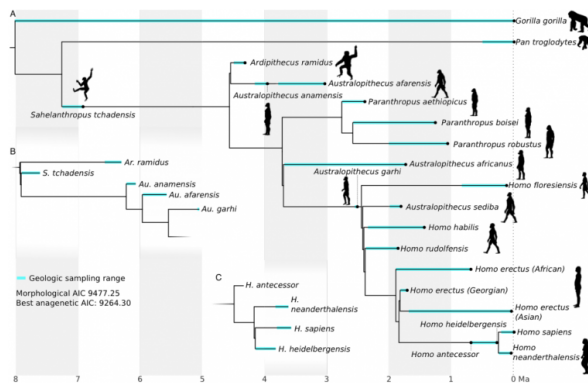


figure 5 - arbre phylogénétique de la lignée humaine d'après parins-fukuchi et coll., 2018 cet arbre représente la meilleure estimation de parins-fukuchi et coll., 2018. à la différence d'autres méthodes de construction d'arbres phylogénétiques, n'utilisant que des données morphologiques, les auteurs estiment, dans l'étude dont est tirée cette figure, qu'il est préférable de recourir en plus aux données de datations des fossiles (lignes bleues) pour proposer des relations d'ancêtres à descendants. c'est ainsi qu'ils identifient *sahelanthropus tchadensis* comme l'ancêtre des hominines ultérieurs, *australopithecus anamensis* comme l'ancêtre d'*australopithecus afarensis*, *australopithecus garhi* comme l'ancêtre du genre *homo*, *homo antecessor* comme l'ancêtre d'*homo heidelbergensis*, qui aurait lui même donné *homo sapiens* et *homo neanderthalensis*. les sous-figures b et c montrent des portions de l'arbre reconstruites sans proposer de relations d'ancêtres à descendants et qui, selon les auteurs, induisent des « résultats erronés ». les silhouettes viennent du site [phylopic](http://phylopic.org).

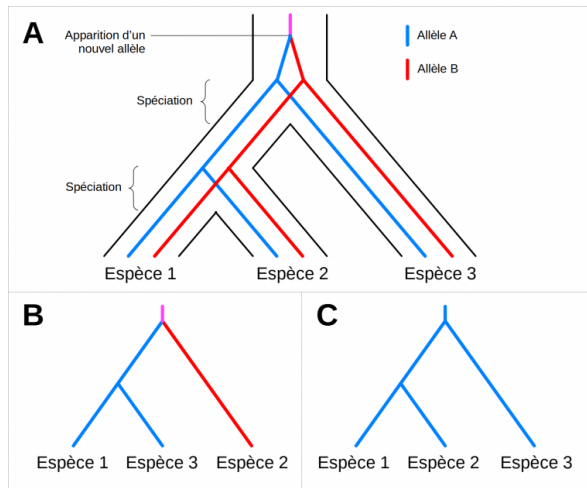
auteur(s)/autrice(s) : parins-fukuchi et coll., 2018 licence : [cc-by-nc-nd](https://creativecommons.org/licenses/by-nc-nd/4.0/) source : [biorxiv](https://doi.org/10.1101/2018.03.01.227881)

trois des quatre spécimens de paranthropes forment un clade, mais la position de sk 835 est plus incertaine car la trichotomie que l'on observe indique que ce nœud n'a pas pu être résolu. autrement dit l'algorithme employé pour construire cet arbre a évalué qu'aucune des trois dichotomies possibles n'était significativement plus vraisemblable que les deux autres.

différents scénarios, cohérents avec chacune de ces trois possibilités, peuvent être imaginés. cependant, avant de surinterpréter ces résultats, il faut se rappeler qu'ils reposent sur assez peu de données. l'ensemble des peptides récoltés chez les paranthropes ne comprend en effet que 17 positions cladistiquement informatives à l'échelle des grands singes, et seulement 2 positions qui sont différentes de ce que l'on rencontre dans le genre *homo*. la première différence se situe en position 637 dans l'alignement de la protéine col17a1 pour sk 835 et sk 14132. ces deux spécimens ont l'état ancestral (proline) qu'ils partagent avec tous les autres grands singes en dehors de la lignée humaine, tandis que le genre *homo* a l'état dérivé (alanine). le peptide en question n'a pas pu être détecté chez les deux autres spécimens. la deuxième différence se situe en position 137 dans l'alignement de la protéine enam. les grands singes africains et le genre *homo* ont ici une arginine tandis que sk 830 et sk 850 ont une glutamine, sk 835 une arginine, et sk 14132 est hétérozygote avec les deux versions du peptide. les orangs-outans présentent aussi une glutamine à la place de l'arginine, ce qui indique une homoplasie, ici une convergence évolutive. il faut donc se garder de toute interprétation hâtive.

le polymorphisme génétique est une autre difficulté importante qu'il faut souligner dans ce genre d'études. le spécimen sk 14132 démontre que plusieurs allèles différents d'un même gène peuvent exister dans une population ancienne, comme c'est le cas dans les espèces modernes. la question du choix de l'allèle pour effectuer la reconstruction phylogénétique se pose alors naturellement. idéalement, les séquences comparées entre deux espèces doivent avoir divergé après la séparation des deux lignées. on dit dans ce cas-ci que les séquences sont orthologues. si les séquences comparées ont commencé à diverger avant la séparation des lignées cela signifie qu'elles descendent d'allèles différents présents chez le dernier ancêtre commun. on dit dans ce cas-là que les séquences sont paralogues. un arbre de gènes ne sera identique à l'arbre des espèces que si l'on compare des allèles orthologues, c'est-à-dire issus d'un même allèle ancestral dont la séquence a divergé entre les espèces après chaque spéciation. si on compare des allèles paralogues, c'est-à-dire des séquences ayant divergé avant les spéciations qu'on cherche à étudier, alors l'arbre des espèces sera biaisé (figure 6).

**figure 6 - le biais de la paralogie dans la reconstruction phylogénétique**  
**a. histoire phylogénétique réelle d'une lignée montrant les évènements de spéciation et de diversification génétique avec l'exemple des allèles a et b issus d'un même allèle ancestral. les allèles a et b sont deux versions d'un même gène qui coexistent chez les trois espèces.**



**b. arbre phylogénétique des espèces biaisé, construit avec les séquences des allèles a des espèces 1 et 3, et la séquence de l'allèle b de l'espèce 2. les allèles a et b sont paralogues.**  
**c. arbre phylogénétique des espèces non biaisé, construit avec les séquences des allèles a des trois espèces. les différentes versions de l'allèle a sont orthologues.**

auteur(s)/autrice(s) : damien  
aubert pennequin licence : [cc-by](https://creativecommons.org/licenses/by/4.0/)

cette paralogie n'est pas facile à identifier, même avec un échantillonnage plus large des populations, car le polymorphisme ancestral peut disparaître chez les descendants. l'un ou l'autre allèle peut alors s'être fixé aléatoirement dans les espèces modernes. ce phénomène porte le nom de tri incomplet des lignées, ou *incomplete lineage sorting* (ils). l'inférence d'un arbre à partir de séquences paralogues reflète alors l'histoire de ces séquences, mais pas nécessairement l'histoire des espèces. l'apparition d'une discordance due à ce biais est d'autant plus probable que les espèces sont apparentées et que les évènements de spéciations se sont succédé rapidement.

ces distorsions peuvent être rectifiées en échantillonnant plusieurs gènes différents comme l'ont fait madupe et ses collaborateurs. dans leur étude, on remarque justement que les arbres phylogénétiques produits à partir des protéines individuelles sont tous différents les uns des autres et sont également différents de celui produit à l'aide de toutes les séquences concaténées. la concaténation des séquences de plusieurs gènes ou protéines permet ainsi de compenser, au moins partiellement, ce type de biais. notons pour finir qu'il existe des méthodes plus complexes, mais statistiquement plus robustes que la simple concaténation, fondées sur la théorie de la coalescence, pour tenir compte de ce phénomène, mais elles n'ont pas été employées dans les travaux présentés ici.

### 3. conclusion

on peut regretter que ces travaux ne soient pas parvenus à récolter davantage de

données moléculaires, comme des peptides plus longs comportant davantage de positions cladistiquement informatives, ou un échantillonnage d'une plus grande diversité de spécimens, comme d'autres espèces de paranthropes ou des australopithèques. les méthodes employées dans cette étude constituent toutefois une véritable percée qui ouvre de nouvelles perspectives pour la recherche en paléanthropologie. il est très probable que les mêmes équipes, et d'autres, soient déjà en train de réaliser de nouvelles investigations afin d'approfondir cette voie prometteuse.

une étude a d'ailleurs permis d'évaluer explicitement le potentiel des protéines anciennes pour l'inférence phylogénétique [17]. les chercheurs ont utilisé les séquences connues d'une grande variété de primates actuels et ont simulé les dégradations typiques qu'elles subissent. ils ont ensuite vérifié que l'arbre phylogénétique reconstruit à partir de ces séquences très parcellaires avait bien une topologie similaire à l'arbre de référence. les résultats montrent que la plupart des familles et des genres sont placés correctement avec des âges simulés d'environ 1-2 ma pour toutes les séquences. de plus, une espèce peut être placée correctement dans l'arbre avec un âge simulé de 5 ma de ses séquences si les séquences des autres espèces sont complètes (ce qui est typiquement ce qu'on cherche à faire quand on veut placer un fossile au sein d'une phylogénie connue d'espèces actuelles). ces résultats sont donc encourageants. par conséquent on peut s'attendre à de nouvelles découvertes dans les années à venir grâce à ces méthodes novatrices.

## CRÉDITS

### AUTEUR(S)/AUTRICE(S)

[damien aubert pennequin](#)

professeur agrégé de sciences de la vie, sciences de la terre et de l'univers (sv-stu), passionné par les sciences de l'évolution, la phylogénétique, la taxonomie, l'algorithmique, la vulgarisation scientifique et la transmission des savoirs.

### RELECTURE SCIENTIFIQUE

[aurélie manin](#)

chercheuse en archéozoologie et archéologie moléculaire, rattachée au laboratoire de la préhistoire à l'actuel : culture, environnement et anthropologie, de l'université de bordeaux.

### MISE EN LIGNE

[pascal combemorel](#)

agrégé de svt, il est le responsable éditorial du site planet-vie depuis septembre 2016.

### LICENCE DU TEXTE DE L'ARTICLE



[creative commons - attribution](#)

---

# bibliographie

1

prüfer, k. ., racimo, f. ., patterson, n. ., jay, f. ., sankararaman, s. ., sawyer, s. ., ...  
pääbo, s. . (2014). the complete genome sequence of a neanderthal from the altai  
mountains. *nature*, 505, 43-49. <http://doi.org/10.1038/nature12886> (original work  
published 2026)

2

krause, j. ., fu, q. ., good, j. m., viola, b. ., shunkov, m. v., derevianko, a. p., &  
pääbo, s. . (2010). the complete mitochondrial dna genome of an unknown hominin  
from southern siberia. *nature*, 464, 894-897. <http://doi.org/10.1038/nature08976>  
(original work published 2026)

3

heyer, e. . (2023). prix nobel de physiologie ou médecine 2022 - svante pääbo,  
précurseur de la paléogénomique. *médecine/sciences*, 39, 181-183.  
<http://doi.org/10.1051/medsci/2023009> (original work published 2026)

4

chen, f. ., welker, f. ., shen, c.-c. ., bailey, s. e., bergmann, i. ., davis, s. ., ... hublin,  
j.-j. . (2019). a late middle pleistocene denisovan mandible from the tibetan plateau.  
*nature*, 569, 409-412. <http://doi.org/10.1038/s41586-019-1139-x> (original work  
published 2026)

5

a  
b

madupe, p. ., koenig, c. ., patramanis, i. ., rüther, p. ., hlazo, n. ., mackie, m. ., ...  
cappellini, e. . (2023). enamel proteins reveal biological sex and genetic variability  
within southern african \textit{paranthropus}. biorxiv.  
<http://doi.org/10.1101/2023.07.03.547326>

6

felsenstein, j. . (1978). cases in which parsimony or compatibility methods will be  
positively misleading. *systematic zoology*, 27, 401-410.  
<http://doi.org/10.2307/2412923> (original work published 2026)

7

felsenstein, j. . (1984). the statistical approach to inferring evolutionary trees and  
what it tells us about parsimony and compatibility. (t. . duncan & t. f. stuessy, éd.).  
new york: columbia university press.

8

hendy, m. d., & penny, d. . (1989). a framework for the quantitative study of  
evolutionary trees. *systematic biology*, 38, 297-309. <http://doi.org/10.2307/2992396>  
(original work published 2026)

9

pardi, f. . (2022). inférence phylogénétique: méthodes basées sur les distances.  
*modèles et méthodes pour l'évolution biologique*, 151-176.

10

gascuel, o. ., & steel, m. . (2006). neighbor-joining revealed. *molecular biology and  
evolution*, 23, 1997-2000. <http://doi.org/10.1093/molbev/msl072> (original work  
published 2026)

11

felsenstein, j. . (1982). numerical methods for inferring evolutionary trees. *the  
quarterly review of biology*, 57, 379-404. consulté à l'adresse  
<http://www.jstor.org/stable/2826886>

12

felsenstein, j. . (2004). *inferring phylogenies*. sunderland, ma: sinauer associates.

13

guéguen, l. . (2022). calculs d'inférence dans les arbres phylogénétiques. *modèles et  
méthodes pour l'évolution biologique*, 177-202.

14

rodrigue, n. . (2022). le paradigme bayésien en phylogénie moléculaire. *modèles et  
méthodes pour l'évolution biologique*, 203-221.

15

parins-fukuchi, c. ., greiner, e. ., maclatchy, l. m., & fisher, d. c. (2018). phylogeny,  
ancestors and anagenesis in the hominin fossil record. biorxiv.  
<http://doi.org/10.1101/434894> (original work published 2026)

16

parins-fukuchi, c. ., greiner, e. ., maclatchy, l. m., & fisher, d. c. (2019). phylogeny, ancestors, and anagenesis in the hominin fossil record. *paleobiology*, 45, 378-393. <http://doi.org/10.1017/pab.2019.12> (original work published 2026)

17

zazueta, r. f., krueger, j. ., alba, d. m., aymerich, x. ., beck, r. m. d., cappellini, e. ., ... lizano, e. . (2024). phylogenetic signal in primate tooth enamel proteins and its relevance for paleoproteomics. *biorxiv*. <http://doi.org/10.1101/2024.02.28.580462> (original work published 2026)