

LE POINT SUR...

CONCISE REVIEW PAPER

L'arbre de la vie a-t-il une structure fractale ?

Is the evolutionary tree a fractal structure?

Jean Chaline^{a*}, Laurent Nottale^b, Pierre Grouc^c

^a Laboratoires de Biogéosciences (UMR CNRS 5561) et Paléobiodiversité et préhistoire de l'EPHE, Université de Bourgogne, Centre des Sciences de la Terre, 6 Bd. Gabriel, 21000 Dijon, France

^b Département d'Astrophysique Extragalactique et Cosmologie (UMR CNRS 8631), Observatoire de Paris, Section de Meudon, 92195 Meudon Cedex, France

^c Université de Versailles-Saint-Quentin-en-Yvelines, 47 Bd. Vauban, 78280 Guyancourt, France

(Reçu le 30 Janvier 1999, accepté après révision le 5 avril 1999)

Abstract — We analyze the time sequences of major evolutionary leaps at various scales, from the scale of the global tree of life, to the scales of orders and families such as sauropod dinosaurs, North American fossil Equidae, rodents, and primates including the Hominidae. In each case we find that these data are consistent with a log-periodic law to a high level of statistical significance. Such a law is characterized by a critical epoch of convergence T_c specific to the lineage under consideration and that can be interpreted as the end of that lineage's capacity to evolve. (© Académie des sciences / Elsevier, Paris.)

tree of life / events / fractals / log-periodic law / spatio-temporal tree structure

Résumé — Les séquences temporelles des grands sauts évolutifs sont analysées à diverses échelles, depuis celle de l'arbre global de la vie jusqu'à celle d'ordres ou familles telles que les dinosaures sauropodes, les équidés fossiles d'Amérique du Nord, les rongeurs et les primates incluant les hominidés. Dans tous les cas, nous trouvons que ces données peuvent être ajustées, avec une haute signification statistique, par une loi log-périodique caractérisée par une époque critique de convergence T_c qui dépend de la lignée considérée et qui peut s'interpréter comme la fin de la capacité d'évolution de cette lignée. (© Académie des sciences / Elsevier, Paris.)

arbre de vie / événements / fractales / loi log-périodique / arborescence spatio-temporelle

Article rédigé à l'invitation du Comité de lecture, sur le conseil de Jean Dorst.

*Correspondence et tirés à part: jchaline@satie.u-bourgogne.fr

C. R. Acad. Sci. Paris, Sciences de la terre et des planètes / *Earth & Planetary Sciences*,
1999. 328, 717-726

Abridged version

Introduction

The jumps between species (Gould et Eldredge, 1977; Chaline *et al.*, 1993) involve bifurcations allowing us to liken the general evolutionary process to a "tree of life" where "branch" lengths represent time intervals between major events. The question raised in this note is whether this tree can be described by a mathematical structure, at least at a statistical level.

By analogy with real trees, we test as a first approximation the simplest possible law, i.e. a self-similar tree. Such a law corresponds to discrete scale-invariance and log-periodic acceleration, and has been suggested as describing various classes of phenomena (notebooks of L. da Vinci published in 1942; Mandelbrot, 1975, 1982; Nottale, 1993; Dubois *et al.*, 1992, Plotnick et McKinney, 1993; Dubois, 1995; Solé *et al.*, 1997, Burlando, 1997, Kirchner et Weil, 1998; Sornette, 1998; Chaline, 1998, 1999).

A fractal tree model

Constructing the evolutionary law and calculating parameters

Let us consider a node in a tree where a branch divides into k sub-branches. Let us assume that the total cross-section before (level n) and after (level $n+1$) the node is preserved. If this section is bidimensional (as for example with conservation of sap flow), this is reflected in the relationship between radii: $k r_{n+1}^2 = r_n^2$. But a more general relationship can be considered by introducing a fractal dimension D : $k r_{n+1}^D = r_n^D$. If we now accept that the tree is fully self-similar (still as a minimum simplifying assumption), the ratio of branch lengths will then be equal to the ratio of their radii, giving $g = k^{1/D}$. As $g > 1$, the total length measured along a given lineage is therefore finite, as it is given by the converging infinite sum: $L_c = L_0 (1 + g^{-1} + g^{-2} + \dots) = g L_0 / (g - 1)$. For a temporal tree, these "lengths" are given by the time interval between two evolutionary events: $L_n = T_{n+1} - T_n$. Convergence of the above series therefore means there is a critical time, T_c , marking the end of the evolutionary process for a given lineage.

If we now take as the time origin the final critical time T_c , we can be sure self-similarity is preserved, because the time interval ratios relative to this origin are always given by g^n (see equation (2) below).

Finally, the law to be tested in what follows can therefore be written:

$$T_n = T_c + (T_0 - T_c) g^{-n} \quad (1)$$

where $g = k^{1/D}$, with k an integer if the number of nodes remains constant at each stage. This law is dependent on two parameters only, g and T_c , which of course have no reason a priori to be constant for the entire tree of life. As regards g , we shall see that a characteristic value nonetheless seems to emerge, but this could be simply a mean value and/or the result of selective optimisation.

Calculating g

From paleontological data, the value of parameter g can be estimated from events taken two by two, once the critical time is known:

$$g = (T_n - T_c) / (T_{n+1} - T_c). \quad (2)$$

Calculating T_c

For a given series with base g , the value of T_c may also be calculated from the two-by-two ratios of dates of the events under consideration (see e.g. Sornette, 1998):

$$T_c(n, p) = (T_n - T_p g^{p-n}) / (1 - g^{p-n}) \quad (3)$$

The stability of the value obtained can be used to test the validity of the law.

The law at different levels of analysis

We test this log-periodic evolutionary law at various levels of analysis: (i) the tree at a global level (Delsemme, 1994), from the first appearance of life to viviparity (*tables 1-2, figures 1-2*); (ii) sauropod dinosaurs (Lambert, 1983; Wilson et Sereno, 1998) (*figure 3*); (iii) rodents (Chaline et Mein, 1979; Gould, 1997; Hartenberger, 1998) (*figure 4*); (iv) the fossil *Equidae* (Chaline 1990) (*table 3, figure 5*); and (v) primates, including the *Hominidae* (Chaline, 1998) (*table 4, figure 6*).

In each case we find that a log-periodic law provides a satisfactory fit for the distribution of dates, with different values of the critical date T_c for different lineages, but with compatible values of the ratio g . The results are found to be of high statistical significance (*figure 7*).

Discussion and conclusion

We end by discussing possible biases and uncertainties in our analysis and conclude that they cannot explain the observed law, which therefore seems to be a genuine one. Analysis of the values of the critical date for the various lineages leads us to interpret it as a limit of the evolutionary capacity of the corresponding species. Such a law does not mean that the role of chance in evolution is reduced, but instead that randomness may occur within a framework which may itself be structured (in a partly statistical way).

1. Introduction

Apparues par des sauts brusques du patrimoine génétique, les espèces biologiques peuvent évoluer graduellement ou rester en stase pendant de longues périodes et s'éteindre individuellement ou à l'occasion de crises d'extinction en masse (Gould et Eldredge, 1977 ; Chaline et al., 1993). Les sauts entre les espèces sont à l'origine de bifurcations qui permettent d'assimiler le schéma évolutif global à un arbre de la biodiversité, "l'arbre de la vie". Ainsi les noeuds de cet arbre peuvent être identifiés aux grands sauts évolutifs, et les "longueurs" des branches aux intervalles de temps entre ces événements majeurs. La question soulevée ici est la suivante : existe-t-il une (ou des) loi(s) statistique(s) pour l'apparition de ces sauts, ou ceux-ci se font-ils d'une manière totalement imprédictible ? En d'autres termes, l'arbre de la vie possède-t-il une structure mathématique ?

Pour tenter de répondre à cette question, on remarquera tout d'abord que les arbres végétaux, qui croissent suivant l'axe spatial vertical, sont caractérisés par des structures auto-similaires de type fractal, une découverte que Mandelbrot (1975, 1982) attribue à Léonard de Vinci [voir ses carnets publiés en 1942]. Nous éprouverons ici l'hypothèse de travail suivant laquelle l'arbre de la vie, comme les arbres végétaux, possède une structure mathématique fractale [voir aussi Nottale (1993)], en la confrontant aux données paléontologiques.

Conscient du caractère extrêmement complexe que devrait revêtir une loi décrivant tous les aspects multiples de l'arbre évolutif (dans l'hypothèse, étudiée ici sans a priori, où une telle loi existe), nous partirons de la loi la plus simple possible, celle d'un arbre auto-similaire et nous la confronterons aux données d'observation. Cette tentative doit être interprétée comme une première approximation, qui n'a pas de raison d'être valide pour tous les sous-ensembles de l'arbre. Nous montrerons en effet que certains des paramètres de la loi obtenue varient suivant les embranchements étudiés (dinosauriens, rongeurs, équidés et primates).

Il faut rappeler que des structures fractales ont déjà été mises en évidence dans les distributions statistiques de séries de spéciation et d'extinction, notamment chez les rongeurs (Dubois et al., 1992), les foraminifères (Plotnick et McKinney, 1993), les ammonites (Solé et al., 1997) et l'ensemble des organismes (Burlando, 1997). Il est vraisemblable que les structures fractales résultent du fait que le vivant est organisé selon une hiérarchie ascendante, une véritable échelle du vivant (Chaline, 1998, 1999). Rappelons également que des lois log-périodiques ont été récemment proposées pour la description de nombreux phénomènes, y compris d'évolution temporelle (Sornette, 1998; Dubois, 1995). Kirchner et Weil (1998) ont critiqué le fait d'interpoler dans le temps les événements, or cet argument avait été discuté et repoussé par Dubois et al. en 1992.

2. Un modèle d'arbre fractal

2.1. Construction de la loi d'évolution et calcul des paramètres

Considérons un point de bifurcation d'un arbre où une branche se divise en k sous-branches. Faisons l'hypothèse de la conservation de la section totale avant (niveau n) et après la bifurcation (niveau $n+1$). Dans le cas où cette section est bidimensionnelle (exemple de la conservation du flux de sève), ceci se traduit par la relation entre rayons: $k r_{n+1}^2 = r_n^2$. Mais une relation plus générale peut être considérée, introduisant une dimension fractale D : $k r_{n+1}^D = r_n^D$. En admettant maintenant qu'il y a une complète auto-similarité de l'arbre (toujours comme hypothèse simplificatrice minimale), le rapport des longueurs des branches sera alors égal au rapport de leurs rayons, soit $g = k^{1/D}$. Comme $g > 1$, la longueur totale mesurée le long d'une lignée donnée est alors finie, car elle est donnée par la somme infinie convergente: $L_c = L_0 (1 + g^{-1} + g^{-2} + \dots) = g L_0 / (g - 1)$. Du point de vue d'un arbre temporel, ces "longueurs" sont données par l'intervalle de temps qui sépare deux événements évolutifs: $L_n = T_{n+1} - T_n$. La convergence de la série ci-dessus signifie alors qu'il existe une époque critique, T_c , qui signe la fin du processus évolutif pour une lignée donnée.

Si l'on prend maintenant comme origine des temps l'époque critique finale T_c , on s'assure que l'auto-similarité est conservée, car les rapports d'intervalles de temps ramenés à cette origine sont toujours donnés par g^n (voir équation 2 ci-dessous).

Finalement, la loi que nous allons éprouver dans la suite s'écrit donc :

$$T_n = T_c + (T_0 - T_c) g^{-n} \quad (1)$$

où $g = k^{1/D}$, avec k entier si le nombre de branchements reste constant à chaque étape. Cette loi ne dépend que de deux paramètres, g et T_c , qui n'ont bien sûr pas de raison a priori d'être constants pour l'ensemble de l'arbre de la vie. En ce qui concerne g , nous

verrons néanmoins qu'une valeur caractéristique semble émerger, mais ceci pourrait n'être qu'une moyenne et/ou le résultat d'une optimisation sélective.

2.2. Calcul de g

A partir des données fournies par la paléontologie, la valeur du paramètre g peut être estimée à partir d'événements pris deux à deux, une fois l'époque critique connue :

$$g = (T_n - T_c) / (T_{n+1} - T_c). \quad (2)$$

2.3. Calcul de T_c

Pour une série à base g donnée, la valeur de T_c peut également être calculée à partir des rapports deux à deux des dates des événements considérés (Sornette, 1998) :

$$T_c(n, p) = (T_n - T_p g^{p-n}) / (1 - g^{p-n}) \quad (3)$$

La stabilité de la valeur obtenue permettra d'éprouver la validité de la loi.

2.4. Analyse statistique

D'une manière plus complète, nous déterminerons simultanément les valeurs de g et T_c à partir des séquences de données observées, par un ajustement numérique (méthode des moindres carrés). On utilisera comme estimateur statistique la variable t de Student associée au coefficient de corrélation dans le diagramme [rang n de l'événement, $\log(T_c - T_n)$]. Pour chaque valeur de T_c , on calcule les valeurs de g et de t . Les valeurs optimisées de g et T_c correspondent alors au pic de la courbe $t(T_c)$. Cependant, les rangs des événements étant a priori classés dans l'ordre chronologique, le niveau de signification statistique associé à une valeur donnée de t n'est pas celui d'une étude de corrélation ordinaire. Des simulations Monte-Carlo ont donc été effectuées pour calibrer cet estimateur en appliquant la même loi à des dates choisies au hasard et triées par ordre chronologique. Les niveaux de probabilité donnés dans la suite sont déduits de cette méthode.

3. La loi aux divers niveaux d'analyse

La difficulté majeure de cette étude réside dans la sélection des événements essentiels. Nous tenterons de définir ou d'utiliser des critères objectifs, les plus indépendants

possibles de tout jugement de valeur, en particulier par le choix de critères physiques (par exemple, changements de niveaux d'organisation: apparition de la vie, des noyaux, de la pluricellularité; acquis fonctionnels : organes de locomotion, homéothermie, organes de préhension...).

3.1. L'arbre global de la vie

Nous allons considérer dans un premier temps l'arbre de la vie, depuis son apparition, à un niveau global, en étudiant l'une de ses "branches" principales menant aux mammifères. Si l'on considère les mammifères vivipares actuels ils ont en commun une structure physique interne qui s'est mise en place à la suite de sauts évolutifs majeurs. La liste des dates de ces sauts est donnée dans le *tableau I*, d'après Delsemme (1994) dont l'étude est antérieure et donc totalement indépendante de la recherche de loi faite ici.

Tableau I. Répartition des événements du transect de l'arbre de vie jusqu'à l'apparition des primates (d'après Delsemme, 1994, complété par la viviparité).

Distribution of events on the evolutionary tree until the appearance of primates [after Delsemme (1994), completed by viviparity].

Evénements	Dates évaluées (Ma)
1 - Apparition de la vie	-3500 ± 400
2 - Eucaryotes (noyaux)	-1750 ± 250
3 - Pluricellularité	-1000 ± 100
4 - Exosquelettes	- 570 ± 30
5 - Tétrapodie	- 380 ± 30
6 - Homéothermie	- 220 ± 20
7 - Viviparité	- 120 ± 20
8 - Primates	- 65 ± 5

Ces dates témoignent d'une accélération log-périodique bien marquée (*figures 1 et 2*). On trouve une époque critique $T_c = -15 (+40, -20)$ Ma, où l'erreur est la demi largeur à mi-hauteur de la courbe t (T_c), qui est ici asymétrique. Sur le même intervalle, on trouve $g = 1.81 \pm 0.07$. La valeur maximale du t de Student est $t = 52$. Ce résultat est statistiquement hautement significatif (probabilité $P < 10^{-3}$ d'obtenir un tel résultat par hasard).

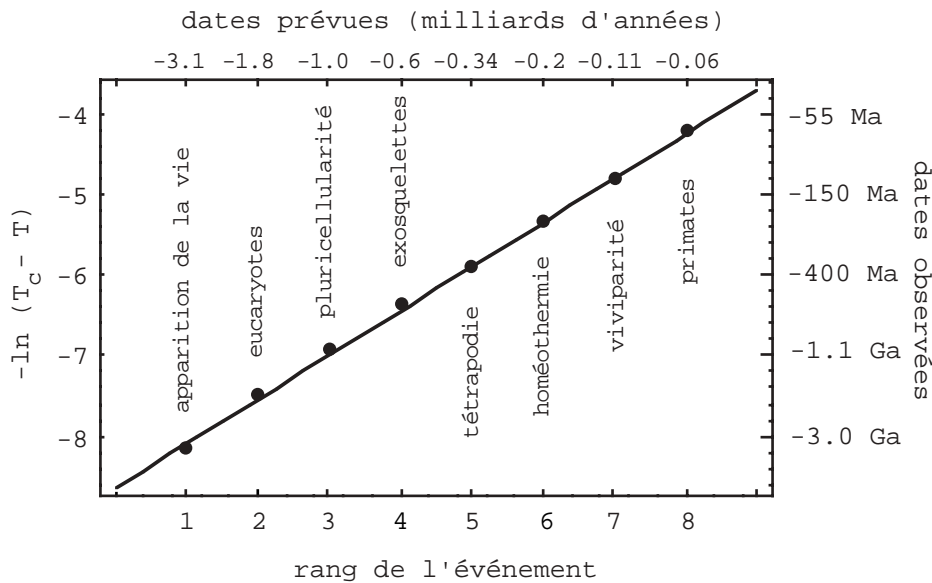


Figure 1. Répartition log-périodique des sauts évolutifs majeurs jusqu'à l'apparition des primates. (On a pris ici $T_c = 2.3$ Ma et $g = 1.73$).

Log-periodic distribution of major evolutionary leaps until the appearance of primates (Here $T_c = 2.3$ Ma and $g = 1.73$).

Pour compléter ces résultats, nous allons maintenant étudier statistiquement la structure plus globale de l'arbre. Notre choix correspond au consensus général de la chronologie des événements novateurs de la vie, depuis ses origines jusqu'aux mammifères (*tableau II*) :

Tableau II. Répartition des événements novateurs de l'histoire de la vie jusqu'à la séparation des Placentaires.

Distribution of innovative evolutionary events until the separation of placental mammals.

Événements	Dates (Ma)
- Apparition de la vie – stromatolithes (= cyanobactéries photosynthétiques) <i>Eobacterium isolatum</i>	- 3400 ± 500
- Accumulation de fer rubané: production d'oxygène	- 2000 ± 200
- Premier eucaryote identifié: acritarches / algues	- 1800 ± 500
- Premier eucaryote multicellulaire identifié (<i>Grypania</i>) ?	- 1100 ± 200
- Disparité des "crown eucaryotes" (Metazoa, Fungi, algues vertes, rouges, brunes, etc...)	- 1000 ± 200
- Diminution du CO2 – glaciations : réduction des bactéries, Cnidaires, diploblastiques	- 600 ± 200
- Disparité des Métazoaires: éponges, arthropodes, mollusques, échinodermes, squelette minéralisé externe, carapace, etc...)	- 550 ± 30
- Apparition du squelette interne: plus ancien vertébré agnathe (Hétérostracé)	- 500 ± 50

- Premiers gnathostomes (Placoderme)	- 410 ± 30
- Premiers tétrapodes pulmonés (<i>Ichthyostega</i>)	- 380 ± 20
- Premier reptile : <i>Westlothiana</i> (oeuf cléidoïque)	- 350 ± 10
- Premier reptile mammalien (homéothermie ?)	- 290 ± 10
- Premier mammifère : (<i>Diarthrognathus-Morganucodon</i>)	- 225 ± 5
- Viviparité (Marsupiaux et Placentaires)	- 120 ± 20

Ces dates sont reportées sur la *figure 2*. Elles sont complétées par:

- trois événements majeurs antérieurs à l'apparition de la vie sur Terre mais essentiels pour les conditions de son apparition: l'origine de l'univers (-16000 ± 2000 Ma), l'enrichissement en éléments "lourds" (C, N, O) du milieu interstellaire ($-11000 / -7000$ Ma) et la formation du Système Solaire (-4600 Ma).

- les dates des grands événements de l'évolution des primates, incluant l'hominisation.

Elles sont alors comparées à celles que l'on peut calculer à partir de la loi $T_n = (2.3 - 16000 / 1.73^n)$ Ma (cf § 3.5). L'accord statistique est remarquable: l'histogramme des écarts δn entre $\log_{1.73}(16000) - \log_{1.73}(T_c - T_n)$ et l'entier le plus proche est fortement piqué autour de zéro, ce qui confirme qu'il apparaît une authentique "quantification". L'écart à une distribution uniforme de cet histogramme est caractérisé par un $\chi^2 = 64.2$ dans 10 intervalles, ce qui correspond à une probabilité $P < 10^{-9}$ d'être issu d'une distribution uniforme.

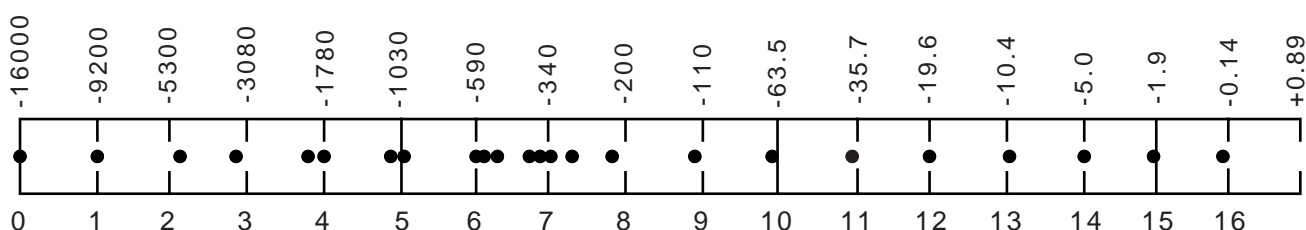


Figure 2. Comparaison entre les valeurs de $\log_{1.73}(T_c - T_n)$ pour les grands événements de l'évolution (y compris l'hominisation et les périodes "cosmologiques") et la distribution des nombres entiers. On a pris $T_c = +2.3$ millions d'années. La probabilité d'obtenir une telle distribution par hasard est inférieure à 10^{-9} .

Comparison of $\log_{1.73}(T_c - T_n)$ values for the major evolutionary events (including hominization and "cosmological" times) and the distributions of integers. Here $T_c = +2.3$ Ma. The probability of obtaining such a distribution by chance is less than 1 in 10^9 .

3.2. Les dinosaures

Nous allons tenter maintenant d'appliquer la même analyse aux dates d'apparition des familles de dinosaures. Si l'on prend en compte toutes les dates relevées par Lambert (1983), on ne détecte aucune accélération. Mais évidemment ces dates constituent un

ensemble hétérogène, qui correspond à un mélange de "noeuds" de l'arbre de rangs différents. On ne peut travailler vraiment que si l'on possède la structure en arbre.

Les données de Wilson et Sereno (1998) recalibrées d'après Odin (1994) sur les sauropodes sont bien adaptées à cette étude, car ils utilisent un critère morphologique rigoureux et ils ont pu reconstituer un arbre évolutif. Dans leur cas l'accélération est très marquée, bien visible sur leur cladogramme. Ils ont identifié 5 événements essentiels (sauropoda, eusauropoda, neosauropoda, titanosauriformes, titanosauria) correspondant à des dates d'apparition que l'on peut estimer à $\{-230, -200, -180, -165, -157\}$ Ma. On trouve pour ces dates (*figure 3*) $T_c = -139$ (-6,+9) Ma et $g = 1.5 \pm 0.1$, résultat hautement significatif ($t = 70$, $P < 2/1000$). En prenant l'échelle de temps choisie par Wilson et Sereno, des dates légèrement différentes peuvent être estimées $\{-230, -208, -182, -167, -156\}$ Ma (voir leur figure 46), qui donnent $T_c = -122$ Ma, $g = 1.33$, et $t = 101$.

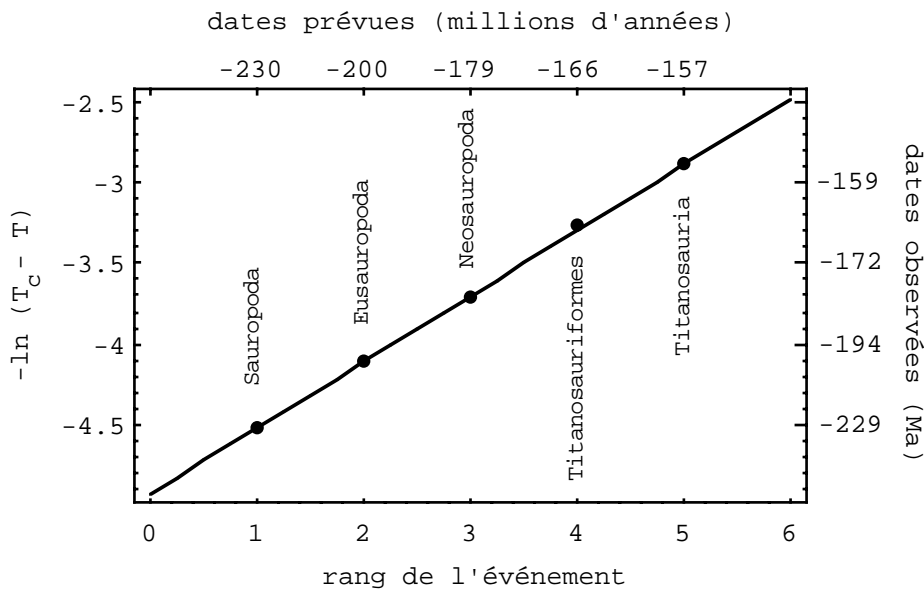


Figure 3. Ajustement des dates des principaux sauts évolutifs des sauropodes par une loi log-périodique définie par $T_c = -139$ Ma et $g = 1.5$.

Fit of dates of main evolutionary leaps for sauropods against a log-periodic law with $T_c = -139$ Ma and $g = 1.5$.

Ce résultat joue un rôle important pour l'interprétation de l'époque critique T_c . En effet, il n'y a plus aucun événement d'apparition de groupes après -156 Ma, beaucoup de ces espèces (Diplodocus, Brachiosaure, Titanosaure), non seulement ne disparaissent pas, mais au contraire restent stables jusqu'à la disparition globale du groupe à -65 Ma. Le T_c semble donc devoir être interprété comme la fin de la capacité évolutive, lorsque la structure se fige et ne peut plus évoluer, et non *a priori* comme une disparition.

3.3. L'ordre des rongeurs

Il était intéressant de considérer le cas d'un groupe possédant une forte biodiversité, pour laquelle on attendrait qualitativement une forte capacité évolutive (Gould, 1997). L'ordre des rongeurs, le plus grand ordre de mammifères puisqu'à lui seul il rassemble la moitié des espèces de ce groupe (Chaline et Mein, 1979), nous fournit un tel cas, pour lequel on s'attend donc à une époque critique plus élevée que dans le cas des primates (le T_C des primates est de +2.3 Ma). Cette prédiction se trouve effectivement vérifiée.

La complexité de l'arborescence des rongeurs, dont la radiation se réalise entre le Paléocène supérieur, il y a plus de 55 Ma, et l'époque actuelle, fera l'objet d'une étude plus complète. Nous nous contenterons dans la présente étude simplifiée, à titre de comparaison avec les autres lignées, d'analyser les pics de la distribution de l'ensemble des dates d'apparition dans l'arbre des rongeurs selon les données de Hartenberger (1998). Une estimation possible de ces dates et de leur rang est : ({-50,1}, {-35,2}, {-26,3}, {-17,4}, {-12,5}, {-7,6},{-2,7}) qui donne un meilleur ajustement avec $T_C = +12 \pm 6$ Ma, $t = 46$ (figure 4, $N = 7$, $P < 10^{-3}$). Un choix un peu différent ({-54,1}, {-43,2}, {-35,3}, {-27,4}, {-18,5}, {-12,6}, {-7,7}, {-2,8}) donne $T_C = +35 \pm 15$ Ma et $t = 80$ ($N = 8$, $P < 2 \cdot 10^{-4}$). Suivant les sous-ensembles considérés de cet ordre, on trouve un T_C grand (variant entre +10 Ma et +60 Ma), toujours supérieur à celui des primates et des équidés américains, en accord avec notre conjecture initiale.

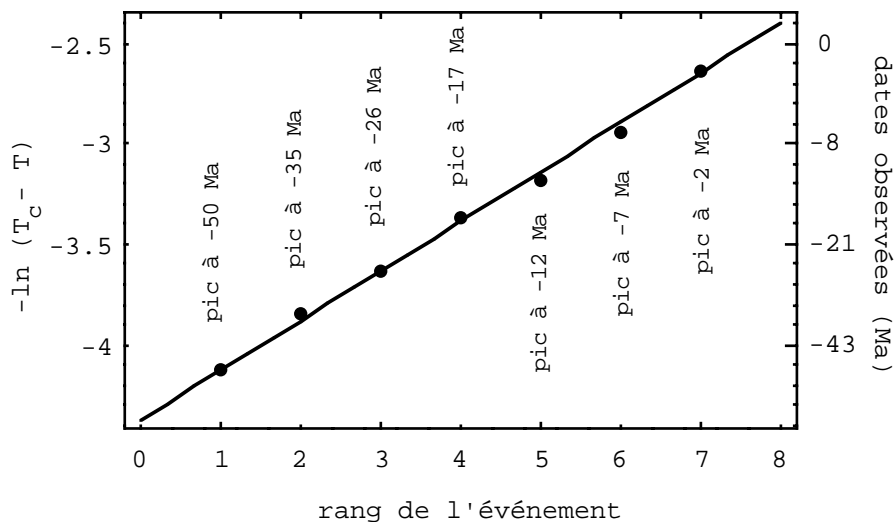


Figure 4. Ajustement des pics de densité dans la distribution des dates d'apparition d'espèces de l'ordre des rongeurs, par une loi log-périodique de $T_C = +12$ Ma et $g = 1.28$.

Fit of density peaks in distribution of rodent species appearances against a log-periodic law with $T_C = +12$ Ma and $g = 1.32$.

3.4. La famille des Equidés fossiles d'Amérique du Nord

L'évolution des équidés fossiles en Amérique du Nord est un exemple qui concerne une diversification partielle au niveau d'une famille de mammifères. On trouve des sauts correspondant à l'apparition des principaux genres aux dates suivantes (*tableau III*) :

Tableau III. Chronologie de l'apparition des genres dans l'histoire des Equidés en Amérique du Nord [selon des données originales de Devillers *in* Chaline 1990].
Evolutionary chronology of North American equid genera [after original data by Devillers *in* Chaline 1990].

Genres	Ages en Ma
<i>Hyracotherium</i>	- 54 ± 5
<i>Orohippus</i>	- 50 ± 5
<i>Epihippus</i>	- 47 ± 5
<i>Mesohippus</i>	- 38 ± 5
<i>Miohippus</i>	- 31 ± 5
<i>Parahippus</i>	- 24 ± 4
<i>Kalobatippus</i> (= <i>Anchitherium</i> archaïque)	- 24 ± 4
<i>Archaeohippus</i>	- 19 ± 3
<i>Merychippus</i>	- 19 ± 3
<i>Hipparion</i>	- 15 ± 3
<i>Protohippus</i>	- 11 ± 2
<i>Nannipus</i>	- 9 ± 2
<i>Plesippus</i>	- 6 ± 2
<i>Equus</i>	- 2 ± 1
Extinction d' <i>Equus</i>	- 0.010

Pour l'analyse de ces données, il nous faut rappeler que notre étude ne porte pas sur les événements en eux-mêmes, mais sur la répartition de leurs dates. Ainsi, aux incertitudes près ($\pm 10\%$), les 3 premiers événements n'en forment qu'un, ce qui est cohérent avec le fait que ces trois premières étapes se réalisent de façon graduelle, sans doute au sein d'un seul et même génome. On vérifie que l'apparition des genres s'effectue selon une loi log-périodique (*figure 5*) : on trouve pour $N = 9$ points (répartis de -54 à -6 Ma sans discontinuité) que le meilleur ajustement correspond à une pente $g = 1.30 \pm 0.05$, et une époque critique $T_c = 0.4 \pm 4$ Ma. Le t de Student associé à cet ajustement est de 39.8 ($P < 10^{-3}$). En gardant l'événement à -2 Ma, on trouve $g = 1.30 \pm 0.03$, $T_c = 0.3 \pm 1$ Ma, $t = 60$ ($P < 10^{-3}$).

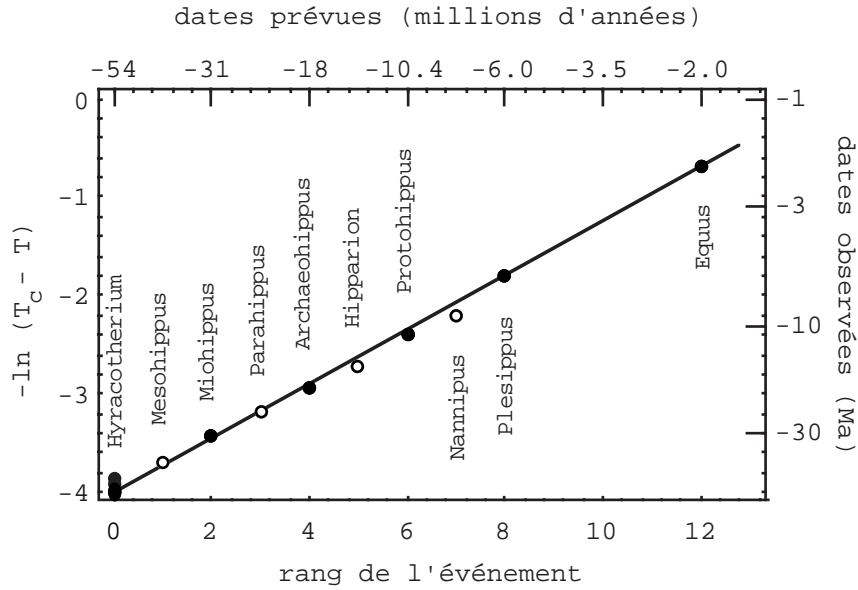


Figure 5. Dates d'apparition des genres d'équidés en Amérique du Nord, comparées à une loi log-périodique avec $T_c = 0$ et $g = 1.32$ (les points de même couleur sont dans un rapport 1.73 : à comparer avec les diagrammes équivalents pour les primates et l'évolution globale).

Appearance dates of North American equid genera compared with a log-periodic law with $T_c = 0$ and $g = 1.32$ (like colored dots are in ratio 1.73 - to be compared with the equivalent diagrams for primates and general evolution).

Il faut remarquer que l'accord n'est possible que s'il manque 3 dates entre -6 et -2 Ma. On prédit donc l'existence de trois événements supplémentaires entre ces deux époques qui n'ont pas été retenus dans le tableau établi à partir des données de Devillers *in* Chaline (1990), à -4.6 , -3.5 et -2.6 Ma, et qui pourraient manquer dans les données paléontologiques actuelles.

Notons en outre que le genre *Equus* disparaît en Amérique du Nord (où les équidés ont été réintroduits par les conquistadors) il y a environ 10.000 ans avec de nombreuses autres espèces de mammifères. Notre résultat est en accord avec cette date d'extinction.

3.5. L'ordre des Primates

Les dates des grands événements de l'histoire des primates (incluant l'homínisation) sont données dans le *tableau IV* (en Ma). Ces dates sont très bien ajustées par une loi log-périodique ($N = 7$, $t = 71.0$, $P = 2 \cdot 10^{-4}$) de $g = 1.736 \pm 0.013$ et de $T_c = 2.25 \pm 0.8$ Ma. Si l'on rajoute les deux événements majeurs que sont l'apparition de mammifères et des placentaires (voir *tableau I*), on trouve un ajustement encore meilleur: $T_c = 1.9 \pm 0.4$ Ma ($N = 9$, $t = 160$, $P \ll 10^{-4}$). La *figure 6* montre la comparaison entre les dates observées et la loi log-périodique correspondante.

Tableau IV. Grands événements de l'histoire des primates. La colonne 3 donne les valeurs de T_c estimées à partir des couples de dates 1 et n. Elles convergent vers 2.3 Ma vers le futur. A partir de cette valeur de T_c , on peut ensuite calculer les valeurs de g à partir des événements pris 2 à 2. Les valeurs correspondantes sont reportées dans la colonne 4. On notera là encore leur grande stabilité: la moyenne est 1.73 ± 0.04 . Cette stabilité signifie que, par exemple, les valeurs des 3 dernières dates peuvent être prédites très précisément à partir des 4 premières, ou encore que la date 4 (qui n'est pas d'origine paléontologique mais est déterminée à partir des estimations de distances génétiques) peut être rétro-prédite par cette loi (prédiction -10.5 Ma) .

Major events in primate evolution. Column 3 shows T_c values estimated from date pairs 1 and n. They converge around 2.3 Ma in the future. Using this T_c value, g values can be calculated from events taken two by two. Corresponding values are shown in column 4. Notice their stability: mean value 1.73 ± 0.04 . This stability means that the values of the last three dates, say, can be predicted very precisely from the first four, or that date 4 (which is not derived from paleontology but from estimated genetic distances) can be reverse-predicted by this law (prediction -10.5 Ma).

Événements	Dates en Ma	$T_c(1,n)$	g
1 - Premier primate	-65 ± 5		
2 - Premier ancêtre anthropoïde	-40 ± 5	+5.8	1.59
3 - Proconsul	-20 ± 2	+2.5	1.90
4 - Ancêtre commun P/G/H	-10 ± 1.5	+3.1	1.81
5 - Australopithèque	-5 ± 1	+2.5	1.68
6 - Premier <i>Homo</i>	-2 ± 0.3	+2.3	1.70
7 - Homme moderne <i>H. sapiens</i>	-0.18 ± 0.02	+2.3	1.72

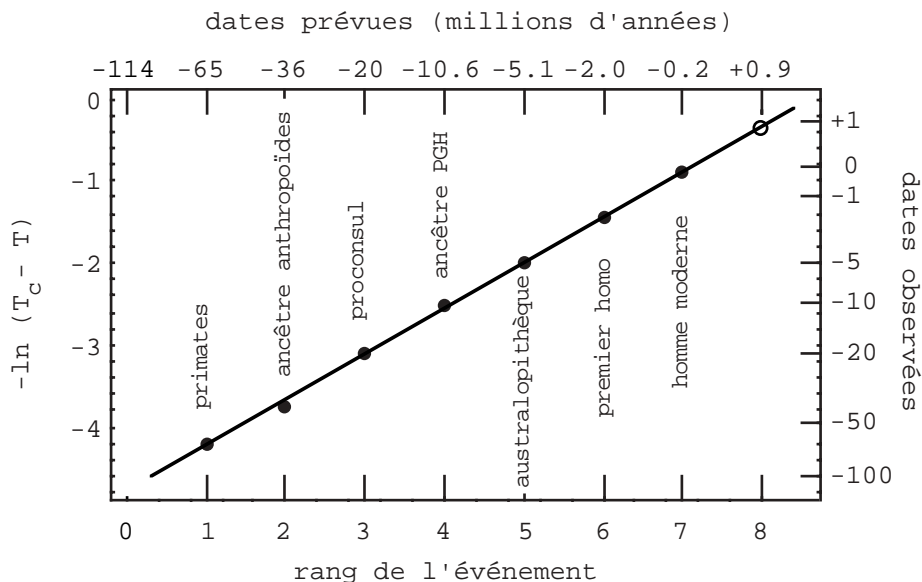


Figure 6. Comparaison entre les dates observées des grands événements de l'histoire des Primates (points) et une loi log-périodique de rapport $g = 1.73$ et d'époque critique $T_c = +2.3$ millions d'années. Le point ouvert correspond à la prédiction issue de cette loi concernant la date du prochain événement évolutif pour cette lignée.

Comparison of dates of major events in primate evolution (dots) and a log-periodic law of ratio $g = 1.73$ and critical time $T_c = +2.3$ Ma. The ring marks the date for the next evolutionary event for the lineage as predicted by the law.

4. Discussion

Les résultats obtenus sont synthétisés dans la *figure 7*. Le caractère partiellement prédictif de notre modèle est évident, à deux niveaux : celui de la structure chronologique de l'arbre de la vie et celui des dates des sauts évolutifs à venir. Avant de développer ce point, il est cependant nécessaire de discuter des biais possibles.

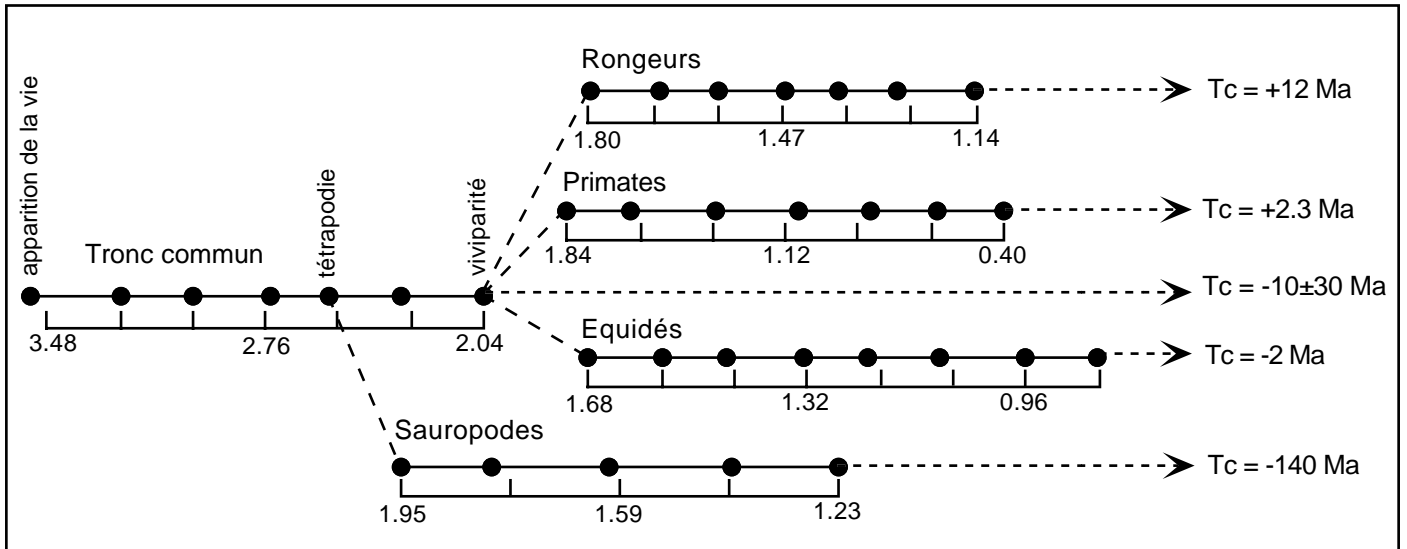


Figure 7. Résumé des principaux résultats de la présente étude. Nous y avons reporté les dates (voir tableaux) des événements marquants de: (i) l'histoire "globale" de la vie ("tronc commun": noter que, strictement, il s'agit déjà d'une branche particulière, la partie "tronc" ne correspondant qu'au premier intervalle, entre l'apparition de la vie et la première ramification en eucaryotes, procaryotes et archaea); (ii) des dinosaures sauropodes, et (iii) des trois lignées de mammifères étudiées ici à titre d'exemple, rongeurs, primates et équidés. Les valeurs des dates sont reportées en échelle logarithmique ramenée à leur date critique, c'est-à-dire $\log(T_c - T)$. L'époque critique T_c ajustée pour chaque lignée est indiquée sur la droite, en millions d'années à partir du présent. La répartition log-périodique (en moyenne) des événements est clairement apparente sur ce diagramme.

Summary of the main findings of this study. We have plotted the dates (see tables) of major events of: (i) general evolution of life ("common trunk": notice that this is strictly already a separate branch, with the "trunk" being just the first interval between the appearance of life and the initial division into eucaryotes, procaryotes and archaea); (ii) sauropod dinosaurs, (iii) the three lineages of mammals studied here as example, i.e. rodents, primates and equids). The numerical values of the dates are given in terms of a logarithmic scale based on their critical date, i.e. $\log(T_c - T)$. The critical time T_c adjusted for each lineage is indicated on the right in millions of years from the present. The (mean) log-periodic distribution of events appears clearly from the diagram.

4.1. Biais de perspective

Le principal problème concernant l'établissement d'une telle loi est le biais qui vient de la difficulté croissante d'obtenir des fossiles à des époques de plus en plus reculées et de l'accroissement de la barre d'erreur sur leur date (effet de "perspective"). Ce biais

pourrait se traduire par une incertitude $\delta T/T$ pratiquement constante sur les dates des événements considérés, ce qui est, effectivement, pratiquement le cas en ce qui concerne la présente étude, et par une densité d'événements constante, non au cours du temps T , mais suivant la variable $\log T$.

Réduire la loi obtenue à un tel biais nous semble cependant impossible. En effet, si la loi n'était que la conséquence de ce biais :

- il faudrait que l'incertitude sur les dates atteigne $\pm 35\%$ (Primates et évolution globale), ce qui semble grandement exagéré par rapport aux 10% généralement adoptés, qui semblent s'accorder avec les variations des estimations suivant les auteurs.
- la loi de log-périodicité ne devrait être qu'une loi moyenne, et l'on ne devrait avoir aucune concentration des points observés autour des points prévus. Or, dans tous les cas, y compris l'évolution "globale", on obtient un regroupement à haute signification statistique autour des dates des événements prévus.
- la valeur de l'époque critique artificielle que l'on calculerait devrait être l'époque actuelle, $T_c = 0$. Ceci est le cas, aux barres d'erreur près (très larges, ± 30 Ma), en ce qui concerne l'évolution "globale" jusqu'à l'homéothermie ($N = 8$, $T < 200$ Ma), ainsi que les Equidés fossiles américains (mais dans ce dernier cas le résultat $T_c = 0$ est une rétro-prédiction vérifiée par les observations). Dans tous les autres cas, la valeur du T_c obtenu est significativement différente de zéro, et montre une haute stabilité autour de sa valeur quels que soient les sous-ensembles d'événements considérés (c'est le cas par exemple du $T_c = +2.3$ Ma pour les primates, cf *tableau IV*).
- le cas des dinosaures sauropodes, où les dates s'échelonnent entre -230 Ma et -156 Ma et convergent vers -139 Ma, exclut quant à lui tout effet de perspective.

Nous concluons que ce biais, bien que jouant sans doute un rôle dans l'analyse des données, ne saurait expliquer cette loi, que nous interprétons donc comme une authentique structuration de l'arbre de la vie.

4.2. Le modèle trichotomique

Une des questions que nous n'avons pas abordées est celle de la valeur des coefficients g obtenus. Dans le cadre d'un modèle d'arborescence régulier, nous avons vu qu'on s'attendait à ce qu'ils soient liés au nombre d'embranchements k et à la dimension D par la relation $g = k^{1/D}$. Or les valeurs obtenues sont toutes compatibles avec $g = 3^{1/2}$ ou sa racine (qui correspond à une densité double de points). Ainsi les données combinées des *tableaux I* et *IV* conduisent à $g = 1.73 \pm 0.03$. Nous sommes donc amenés à faire la conjecture qu'au moins certains sous-ensembles de l'arbre se caractériseraient par un nombre d'embranchements moyen $k = 3$ et une dimension $D = 2$.

Inversement il semble probable que certaines explosions évolutives comme celle du Cambrien ont donné un nombre d'embranchements supérieurs, mais seuls comptent finalement ceux qui restent après l'action de la sélection naturelle.

On peut également faire la conjecture que la valeur de D soit liée à la partie spatiale de l'arbre et donc que $D = 2$ manifeste le caractère essentiellement bi-dimensionnel de la biosphère. Ceci pourrait être éprouvé par des études futures de lignées moins soumises à cette contrainte (par exemple, répartition localement tridimensionnelle).

4.3. Validité des prédictions

On peut enfin utiliser la loi obtenue pour prédire l'époque du prochain saut majeur de même ampleur que l'apparition de l'homme moderne (*Homo sapiens*). On le trouve à ≈ 0.8 Ma vers le futur, ce qui fixerait à environ 1 Ma la période totale où l'homme moderne représente la phase la "plus" évoluée de la biosphère. Quelle valeur attribuer à une telle prédiction ? On peut évidemment s'interroger sur les perturbations que risque d'apporter à cette loi (si sa validité se confirmait) le fait même de la connaître... En outre l'homme devenant maître de son évolution risque de la modifier par son action. Il est néanmoins remarquable que l'homme moderne semble apparaître dans la loi obtenue comme un noeud intermédiaire et transitoire, mais a priori pas comme une limite.

Inversement, il faut aussi prendre en compte les effets de taille finie, comme dans le cas par exemple de la prédiction par des lois log-périodiques des cracks boursiers ou des tremblements de terre [voir Sornette (1998)], où l'époque de divergence théorique ne peut être qu'une limite supérieure pour l'époque de crise réelle. Un problème ouvert est donc la détermination du seuil, soit sur les valeurs d'une variable divergente (indice boursier pour les cracks, concentration d'ions pour les tremblements de terre...), soit sur les écarts de plus en plus petits entre crises secondaires, à partir duquel on entre réellement dans la crise finale. C'est ainsi que, à l'échelle de l'intervalle total d'évolution de la vie sur Terre, soit environ 3.5 milliards d'années, il n'est pas du tout impossible qu'une séparation entre grands sauts évolutifs qui ne serait plus que de 1 million d'années seulement corresponde au dépassement d'un tel seuil. Dans ce cas, l'apparition de l'homme moderne pourrait être considérée comme la crise finale pour notre lignée : mais il n'y a pas de raison qu'il en soit de même pour l'ensemble de l'arbre, comme l'indique par exemple l'obtention d'un T_c plus éloigné dans le futur pour les rongeurs.

Cette loi ne doit pas être interprétée de façon finaliste. L'existence d'une telle loi ne peut s'expliquer, à notre avis, que par l'existence de structures de l'arbre évolutif pris globalement. C'est en poursuivant cette étude par une analyse détaillée de l'ensemble de l'arbre que le sens des lois obtenues pourra apparaître clairement.

Il faut également se garder de mal ou trop hâtivement interpréter la signification de l'époque T_c (et donc de pousser trop loin l'analogie avec les vrais arbres). Il s'agit de la

date d'une "crise" majeure pour la lignée considérée et d'après les données des dinosaures, il pourrait s'agir de la perte des potentialités évolutives, ce qui ne signifie pas extinction immédiate. De plus, la fin d'une loi n'est pas la fin de toute loi, mais peut être la transition vers une loi nouvelle. Il faut insister également sur le fait que nulle part, dans la détermination de la loi obtenue, la nature des événements n'est intervenue : seules leurs dates ont été prises en compte.

Si cette loi empirique est confirmée ultérieurement, une conséquence potentiellement importante doit être soulignée : la loi implique que les divers groupes zoologiques, ayant un T_c différent, ont des durées d'existence distinctes et limitées. Il est clair que s'il faut se garder d'interpréter systématiquement l'époque critique comme celle d'une extinction définitive, il n'en reste pas moins vrai que cette propriété de convergence temporelle fournit un cadre nouveau de compréhension de certaines extinctions de groupes (en dehors des accidents conduisant à des extinctions massives d'origine environnementale). Ceci a été illustré par notre étude des équidés fossiles d'Amérique du Nord, effectivement disparus à la date critique (aux barres d'erreur près), prédite à partir de dates beaucoup plus anciennes. Il sera intéressant d'appliquer ce type d'analyse à d'autres groupes, ce que nous ferons dans un travail futur.

5. Conclusion

Les résultats obtenus nous ont permis de valider notre hypothèse de travail initiale, selon laquelle la représentation d'un arbre fractal peut être transposée dans le domaine temporel. L'arbre de vie dans tous ses détails pourrait se révéler plus complexe qu'un arbre végétal. Mais le point essentiel développé ici est l'idée qu'une description quantitative en terme de structure fractale est possible, à condition d'aller au delà de la self-similarité stricte (Nottale, 1997).

La loi obtenue semble impliquer l'existence d'un déterminisme sous-jacent à l'évolution au niveau structurel, qui ne remet pas en compte son indéterminisme profond dû à la contingence (Gould, 1997), sachant qu'elle n'apporte aucune prévision de la nature des événements. C'est une loi qui possède néanmoins un certain niveau de prédictibilité au niveau chronologique et qui peut donc être testée, même si les prédictions dans le domaine de la paléontologie relèvent du très long terme.

Remerciements. Nous remercions J.-L. Hartenberger pour les données inédites qu'il a bien voulu nous communiquer, Y. Bouligand, F. Cézilly, B. David, J. Dercourt, J.L. Dommergues, J. Dubois, F. Magniez, S. Montuire, M. Remoissenet, D. Sornette et E. Verrecchia pour leurs critiques et de très utiles discussions.

RÉFÉRENCES

- Burlando B. 1997. The fractal geometry of evolution, *J. Theor. Biol.*, 163, 161-172.
- Chaline J. 1990. *Paleontology of Vertebrates*, Springer-Verlag, New-York, 186p.
- Chaline J. 1998. Vers une approche globale de l'évolution des Hominidés, Le Point sur..., *C. R. Acad. Sci. Paris*, 326 (3II), 307-318.
- Chaline J. 1999. *Les Horloges du vivant*, Hachette, Paris, 236p.
- Chaline J. et Mein P. 1979. *Les rongeurs et l'évolution*, Doin, Paris, 235p.
- Chaline J., Laurin B., Brunet-Lecomte P. et Viriot L. 1993. Morphological trends and rates of evolution in Arvicolids (Arvicolidae, Rodentia) at species level : toward a Punctuated Equilibria/Disequilibria Model, in : Chaline J. et Werdelin L. (éds.), *Modes and tempo of Evolution in the Quaternary*, Quaternary International, 19, 27-39.
- Delsemme A. 1994. *Les origines cosmiques de la vie*, Flammarion, Paris, 382p.
- Dubois J. 1995. *La dynamique non linéaire en physique du globe*, Masson, Paris, 265p.
- Dubois J., Chaline J. et Brunet-Lecomte P. 1992. Spéciation, extinction et attracteurs étranges, *C. R. Acad. Sci. Paris*, 315(II), 1827-1833.
- Gould S.J. 1997. *L'éventail du vivant*, Le Seuil, Paris, 304p.
- Gould S.J. et Eldredge N. 1977. Punctuated equilibria: the tempo and mode of evolution reconsidered, *Paleobiology*, 3(2), 115-151.
- Hartenberger J.-L. 1998. Description de la radiation des Rodentia (Mammalia) du Paléocène supérieur au Miocène ; incidences phylogénétiques, *C. R. Acad. Sci. Paris*, 316(II), 439-444.
- Kirchner J.W. et Weil A. 1998. No fractals in extinction statistics, *Nature*, 395, 337-338.
- Lambert D. 1983. *Collins Guide to Dinosaures*, Collins, Hong-Kong, 256p.
- Mandelbrot B.B. 1975. *Les Objets Fractals*, Flammarion, Paris, 203p.
- Mandelbrot B.B. 1982. *The Fractal Geometry of Nature*, Freeman, San Francisco, 460p.
- Nottale L. 1993. *Fractal Space-Time and Microphysics: Towards a Theory of Scale Relativity*, World Scientific, Singapour, 333p.
- Nottale L. 1997. Scale Relativity, in *Scale Invariance and Beyond*, eds. B. Dubrulle, F. Graner & D. Sornette, EDP Sciences / Springer, 249-261.
- Odin S. 1994. Geological Time Scale, Le Point sur..., *C. R. Acad. Sci. Paris*, 318 II(1), 59-71.
- Plotnick R.E. et McKinney M. 1993. Evidence of self-organization in planktic foraminiferal evolution: implications for interconnectedness of palaeosystems, *Palaios*, 8, 202-212.

- Solé R.V., Manrubia S.C., Benton M. et Bak P. 1997. Self-similarity of extinction statistics in the fossil record, *Nature*, 388, 764-767.
- Sornette D. 1998. Discrete scale invariance and complex dimensions, *Physics Reports*, 297, 239-270.
- Vinci L. de. 1942. *Les Carnets de Léonard de Vinci, 1. Botanique*, I, Gallimard, Paris, 323.
- Wilson J.A. et Sereno P.C. 1998. Early Evolution and Higher-level Phylogeny of Sauropod Dinosaurs, *Journal of Vertebrate Paleontology*, 2(18), 1-68.