

# QUE REPRÉSENTE-T-ON AVEC UN ARBRE ?

Guillaume LECOINTRE

UMR 7138

Département Systématique et Évolution  
Muséum national d'Histoire naturelle  
CP 39, 57 rue Cuvier, 75005 Paris.

lecointr@mnhn.fr

*Qui sait à quel instant de la succession de ces générations animales nous en sommes ? Qui sait si ce bipède déformé, qui n'a que quatre pieds de hauteur, qu'on appelle encore dans le voisinage du pôle un homme, et qui ne tarderait pas à perdre ce nom en se déformant un peu davantage, n'est pas l'image d'une espèce qui passe ? Qui sait s'il n'en est pas ainsi de toutes les espèces d'animaux ?* (Denis DIDEROT, *Le Rêve de d'Alembert*, 1769.)

## INTERROGER LA MÉTAPHORE DE L'ARBRE

La figure de l'arbre comme forme d'expression d'un ordre dans le vivant et, le plus souvent, d'un déploiement temporel de celui-ci est apparue à la charnière des XVIII<sup>e</sup> et XIX<sup>e</sup> siècles. Le problème de l'arbre, précisément, c'est qu'il est une métaphore qui, au cours de l'histoire des sciences, a été utilisée tantôt comme métaphore pédagogique à vocation classificatoire sans transformisme (chez A. AUGIER) ou bien avec transformisme généralisé (J.B. DE LAMARCK, C.H. DE BARBANÇOIS), tantôt comme figure généalogique (chez G. BUFFON, C. DARWIN), tantôt comme échelle des êtres devenue évolutionniste (chez E. HAECKEL, A. ROMER), tantôt comme phylogénie au sens moderne (chez W. HENNIG). L'usage récurrent et multiple de la métaphore de l'arbre explique bien des malentendus passés et actuels, et au premier

chef le malentendu entre HAECKEL et DARWIN sur le mot «phylogénie». Pour y voir clair dans les débats actuels, il faut distinguer dans la métaphore de l'arbre trois entités sous-jacentes qu'elle a servi tour à tour à désigner : la généalogie théorique, le graphe connexe, et enfin, la phylogénie.

Ces trois entités résident dans des registres cognitifs distincts. La généalogie théorique, comme son nom l'indique, fixe le cadre théorique des relations que doivent entretenir les organismes (ou les entités sujettes à transmission) dans leur perpétuation. Les lignes de descendance sont tracées entre des entités abstraites (A, B, C, D...) et l'on trouve ce genre de figures dans DARWIN (1859) ou dans HENNIG (1950). La forme générale de la généalogie théorique dépend de nos connaissances des processus de transformation à l'œuvre. Le graphe connexe, lui, est du registre technique, le registre des moyens par lesquels on représente les résultats concrets d'une investigation ponctuelle sur des organismes concrets. L'un des pièges classiques de la reconstruction phylogénétique appliquée aux espèces biologiques, sur lequel nous reviendrons, est que le graphe en question, tel qu'il est produit majoritairement aujourd'hui, ne reflète en première instance ni des liens généalogiques entre entités concrètes, ni même des degrés relatifs d'apparementement entre entités concrètes, mais des degrés relatifs

de partage d'attributs. Enfin, le troisième niveau, épistémologique, est celui de la phylogénie : il est la résultante interprétative du graphe (niveau 2) à la lumière de son cadre théorique (niveau 1) ; où l'on synthétise les degrés relatifs d'apparement et les homologues qui les soutiennent. Lorsque la phylogénie des gènes suit assurément celle des espèces qui les portent, il y a une superposition directe entre les niveaux 2 et 3. Un graphe issu d'un calcul nous fournit directement « la » phylogénie des espèces, dit-on. Mais lorsque ce n'est pas le cas, notamment chez les micro-organismes, un graphe connexe non cyclique ne suffit plus à lui tout seul à résumer une interprétation complète des degrés relatifs de partage et d'apparement entre entités. C'est là que cette distinction devient utile. Enfin, la phase d'interprétation se poursuivant vers une phase de synthèse, on pourrait même trouver un quatrième sens à la métaphore de l'arbre, la phylogénie arrivée à son étape synthétique. Dans tous les musées d'aujourd'hui (ou presque) on dessine des arbres qui ne sont pas directement tirés d'un calcul à partir d'une matrice, mais qui sont des synthèses des connaissances sur les relations d'apparement entre entités concrètes. Ainsi, lorsqu'on parle de l'« arbre de la vie », il ne s'agit pas d'un graphe calculé à partir d'une matrice de caractères, mais le plus souvent d'une synthèse (niveau 4) si les entités sont concrètes, parfois d'un cadre théorique lorsque les entités sont abstraites (niveau 1). Résumons-nous. Au premier niveau, l'arbre généalogique dessine la forme théorique des lignes de transmissions entre entités abstraites ; il n'est pas tiré d'un calcul. Au second niveau, le graphe qu'on appelle « arbre » dessine des partages d'attributs entre entités concrètes et il est bien tiré d'un calcul. Au troisième niveau l'arbre phylogénétique synthétise les deux niveaux précédents en interprétant les degrés relatifs d'apparement et les homologues secondaires entre entités concrètes ; il peut être tiré d'un calcul ou de plusieurs calculs successifs (forêts d'arbres, arbres consensus divers...). Enfin, dans le prolongement du précédent, le niveau de la synthèse des connaissances dessine des « arbre de la vie » qui ne sont pas directement tirés d'un calcul et qui concernent des entités concrètes. Pour

comprendre la nécessité de ces distinctions entre les sens cachés du mot « arbre », il faut passer d'abord par l'histoire des arbres et de leurs significations. C'est ce que nous allons faire à présent.

### Naissance de la filiation des espèces

L'intuition selon laquelle les espèces sont généalogiquement affiliées entre elles diffuse dans la seconde moitié du XVIII<sup>e</sup> siècle. On a coutume d'associer spontanément cette intuition au transformisme, mais sur ce point il faut rester prudent. Par exemple, chez le précurseur du transformisme qu'est Benoist DE MAILLET (1755), les origines des êtres terrestres à partir des êtres marins restent distinctes les unes des autres, réalisant une sorte de transformisme sans généalogie possible des espèces. De même, une véritable filiation des espèces entre elles est logiquement impliquée par la physique de la transformation des espèces de Pierre-Louis MOREAU DE MAUPERTUIS en 1750, mais ne restera jamais qu'implicite. Erasm Darwin fut beaucoup plus explicite dans *Zoonomie ou Lois de la vie organique* en 1794 :

*Serait-ce [...] une témérité d'imaginer que, dans la longue suite des siècles écoulés depuis la création du monde, peut-être plusieurs millions de siècles avant l'histoire du genre humain, tous les animaux à sang chaud sont provenus d'un filament vivant que la grande cause première a doué de l'animalité, avec la faculté d'acquiescer de nouvelles parties accompagnées de nouveaux penchants dirigés par des irritations, des sensations, et ainsi posséder la faculté de continuer à se perfectionner par sa propre activité inhérente et de transmettre ces perfectionnements de génération en génération, à sa postérité et dans les siècles des siècles ?*  
(DARWIN, 1794).

On trouve ici, chez le grand-père de Charles, l'idée de filiation des êtres inscrite dans un temps long, et la transformation des espèces avec transmission des modifications à la descendance. Le travail d'Erasm DARWIN manquait cependant de base observationnelle. Qui dit filiation dit généalogie. La généalogie est une figure représentant des lignes entre des parents et des descendants identifiés. Si cette généalogie est évidente au sein de l'espèce comme représentant légitimement les

liens familiaux entre individus, elle devient plus abstraite s'il s'agit de représenter des liens d'ancêtres à descendants entre populations, sous-espèces, et marque une révolution conceptuelle si ces liens sont représentés entre espèces ou entre entités taxonomiques de plus haut rang. C'est l'horticulteur Antoine Nicolas Duchesne qui franchit le pas le premier en 1766 en publiant une généalogie des fraisiers (DUCHESNE, 1766). Georges Buffon publia bien une carte généalogique des chiens en 1755 (BUFFON, 1755), mais resta au seuil d'une véritable généalogie des espèces, ballotté qu'il était entre une conception de la «dégénération» des espèces et une négation de celle-ci justifiée par un recours à la révélation. Cependant, en 1766, la métaphore généalogique appliquée aux espèces semble poindre chez lui et prendre le dessus :

*Une considération très importante et dont la vue est très étendue, c'est celle du changement des espèces-mêmes, c'est cette dégénération plus ancienne et de tout temps immémoriale qui paraît s'être faite dans chaque famille ou, si l'on veut, dans chacun des genres sous lesquels on peut comprendre les espèces voisines et peu différentes entre elles ; nous n'avons dans tous les animaux terrestres que quelques espèces isolées, qui, comme celle de l'Homme, fassent en même temps espèce et genre ; l'éléphant, le rhinocéros, l'hippopotame, la girafe forment des genres ou des espèces simples qui ne se propagent qu'en ligne directe et n'ont aucune branche collatérale ; toutes les autres paraissent former des familles dans lesquelles on remarque ordinairement une souche principale et commune, de laquelle semblent être sorties des tiges différentes et d'autant plus nombreuses que les individus dans chaque espèce sont plus petits et plus féconds. [...] Sous ce point de vue, le cheval, le zèbre et l'âne sont tous trois de la même famille ; si le cheval est la souche ou le tronc principal, le zèbre et l'âne seront des tiges collatérales ; le nombre de leurs ressemblances entre eux étant infiniment plus grand que celui de leurs différences, on peut les regarder comme ne faisant qu'un même genre, dont les principaux caractères sont clairement énoncés et communs à tous trois (BUFFON, 1766).*

On remarque dans ces écrits que si la généalogie des espèces n'est pas dessinée, elle est écrite. Les mots «branche», «souche», «tronc», «tige»

renvoient métaphoriquement à *un arbre* dont on sait qu'il est généalogique, puisqu'on parle de «changement des espèces-mêmes», de «ligne directe» et de «famille». Ici le mot *famille* relève de la métaphore généalogique et non du rang taxonomique : Buffon ne parle dans ce passage que de la diversification des espèces au sein de genres. D'autre part, il y a adéquation directe entre le déploiement généalogique des espèces par «dégénération» et la classification puisque qu'on va ranger dans un même «genre» les espèces partageant la même «souche». Le pas est donc franchi... mais Buffon n'ira pas plus loin par écrit. Il laissera à d'autres le soin de généraliser son transformisme qui n'est que limité. L'inscription de la figure de l'arbre dans une dimension temporelle (non pas absolue mais relative) s'effectue dès 1801 chez le botaniste français Augustin AUGIER (1801). Mais l'arbre implique-t-il chez ces auteurs une référence à la généalogie des espèces ? Chez Duchesne, un transformisme est admis. C'est beaucoup moins clair chez Augier et Buffon. Même s'il exprime par son arbre «un ordre que la nature paraît avoir suivi dans le règne végétal», Augier n'est pas transformiste et s'en réfère à l'ordre du créateur. L'arbre n'est chez lui qu'une structure classificatoire, non reliée à un message transformiste. La figure de l'arbre est véritablement associée à une généalogie des espèces à grande échelle en 1809 chez Jean-Baptiste LAMARCK. En 1816, Charles Héliou de BARBANÇOIS-VILLEGONGIS, agronome, publie un arbre d'esprit lamarckien encore plus précis et utilise explicitement le terme de «filiation des animaux», laquelle guide la classification. Mais chez Barbançois, les transmutations d'espèces dont il fait le récit, à l'appui de sa figure arborescente, ne sont pas fondées sur l'analyse raisonnée des caractères et, sans doute pour cette raison, son arbre rencontrera peu d'échos (cf. TASSY, 1991). L'arbre est représenté dès 1837 dans les *notebooks* de Charles Darwin, en 1856 chez Alfred Russel WALLACE (1856). L'arbre qui s'impose, surtout dans le dernier tiers du XIX<sup>e</sup> siècle, est appelé «phylogénie» en 1866 par Ernst HAECKEL. La figure d'arbre la plus connue est, en effet, certainement celle que publia Ernst HAECKEL en 1874. DARWIN (1859) ne parle que

de généalogie et dessine un modèle de généalogie. Dans un premier temps, il ne dispose pas du mot « phylogénie ». HAECKEL (1866) ayant lu Darwin, il a bien compris que les meilleures classifications doivent être généalogiques. Mais il crée le terme de « phylogénie » pour parler de l'enchaînement des espèces en une « histoire du développement paléontologique des espèces organiques ». Cette histoire est pour Haeckel l'*histoire* de l'évolution, l'évolution étant elle-même un concept ontogénétique, relevant du développement des individus. Mais DAYRAT (2003) montre que si Haeckel a pour maîtres à penser Lamarck, Goethe et Darwin, sa phylogénie tient plus de Goethe et de Lamarck que de Darwin. En effet, la phylogénie de Haeckel ne montre des successions, des enchaînements que sur son tronc, ses branches latérales étant dépourvues de manifestations. Non seulement le tronc est le seul segment qui vaille qu'on y montre quelque chose, mais ce qui s'y passe n'est pas une généalogie d'espèces, mais une succession de morphologies (parfois celle d'un seul organe) par lesquelles une « chaîne de progéniteurs » est passée. Si l'on retire les branches latérales qui ne font office que d'ornements, l'arbre de Haeckel n'est qu'une série linéaire d'étapes morphologiques majeures idéales le long de la descendance menant à une espèce donnée, donc sur le tronc menant à cette espèce. Les branches latérales n'étant qu'illustratives, l'arbre de Haeckel tient plus du scalisme idéaliste que d'une véritable généalogie. Cependant, Darwin, en retour et sur le tard, va reprendre le mot « phylogénie » pour parler de la grande généalogie des êtres organisés (DARLU & TASSY, 1993 ; DAYRAT 2003, 2005). Aucune de ces deux définitions du mot « phylogénie » ne sont celles qui sont retenues aujourd'hui.

Autre figure majeure de l'arbre est sans doute celle que publia DARWIN dans *L'Origine des espèces* en 1859. Cette figure est à usages multiples et demande qu'on s'y arrête (voir section « Interroger l'arbre généalogique théorique »). Mais avant cela, il faut aborder la difficulté de traduction du mot « *descent* », qui a sans doute participé aux difficultés de compréhension du texte de Darwin par les francophones.

## Difficultés de traduction

Pour comprendre de quoi Darwin parle exactement lorsqu'il parle de « *descent with modification* », il faut s'en référer au passage suivant :

[...] *the natural system is founded on descent with modification ; that the characters which naturalists consider as showing true affinity between any two or more species, are those which have been inherited from a common parent, and, in so far, all true classification is genealogical ; that community of descent is the hidden bond which naturalists have been unconsciously seeking, and not some unknown plan of creation, or the enunciation of general propositions, and the mere putting together and separating objects more or less alike.*

Le public francophone parle de « descendance avec modification » comme le guide de la classification. Ceci provient probablement de ces « faux amis » entre l'anglais et le français. En effet, les Anglais utilisent « *descent* » tout aussi bien pour ascendance ou descendance, voire même filiation ou extraction. Par exemple, dans le *Longman contemporary English*, on trouve : « *Descent : family origins of the stated type : She is of German descent.* » Dans ce sens, « *descent* » se traduit en français par la métaphore de l'extraction généalogique : elle est d'extraction allemande, elle est de souche allemande (sur ce point, voir TORT, 1996, tome 1, p. 1189-1191). Voici comment Edmond Barbier traduit le passage mentionné ci-dessus (DARWIN, 1880) :

[...] *que le système naturel a pour base la descendance avec modification, et que les caractères regardés par les naturalistes comme indiquant des affinités réelles entre deux ou plusieurs espèces sont ceux qu'elles doivent par hérédité à un parent commun. Toute classification vraie est donc généalogique ; la communauté de descendance est le lien caché que les naturalistes ont, sans en avoir conscience, toujours recherché, sous prétexte de découvrir, soit quelque plan inconnu de création, soit d'énoncer des propositions générales, ou de réunir des choses semblables et de séparer des choses différentes.*

Mais à y regarder de plus près, ce texte ne se comprend que si l'on traduit le mot « *descent* » de la seconde phrase par « ascendance » : en effet,

«ceux [les caractères] qui ont été hérités d'un parent commun» seraient-ils des signes de la descendance commune? On n'hérite pas de sa descendance! Peuvent-ils guider la classification, laquelle se fonde sur la communauté d'ascendance (et non pas la communauté de descendance)? Le naturaliste qui doit classer des espèces ne dispose pas de leur descendance! Cette traduction installe donc une confusion chez les francophones. Pire, cette confusion mélange deux approches scientifiques courantes des sciences naturelles : les «*patterns*» (science des patrons) et les «*processes*» (science des processus). Si l'on travaille en sciences des processus, c'est-à-dire que l'on met en évidence expérimentalement des relations de cause à effet, le «*descent with modification*» est le processus par lequel les espèces dérivent les unes des autres et se transforment au cours du temps. Il se déroule dans le temps biologique, du passé vers le présent ou du présent vers le futur. La science des processus, sachant la cause, prédit l'effet. Lorsqu'un naturaliste possède en mains des espèces à classer, ce n'est pas cette science-là qu'il peut mobiliser. Le naturaliste classificateur possède les effets en mains (le partage des attributs entre espèces) et doit en inférer la cause (ascendance commune ou convergence?). Si l'on se place dans les sciences des patrons, au sein desquelles il s'agit de structurer et d'interpréter l'agencement des entités de la nature, alors Darwin nous dit que le partage des caractères doit être parié, en première instance, comme étant dû à l'ascendance commune (et non à la descendance commune!) et que cet agencement doit être guidé par les degrés d'affiliation résultant du passé. Sachant les effets (le partage des caractères), le classificateur en infère la cause (tel caractère est obtenu chez ces espèces par ascendance commune, tel autre caractère est obtenu chez telles autres espèces par convergence). La généalogie du vivant doit donc être considérée par le naturaliste à rebours du processus biologique, si le naturaliste veut parvenir à ses fins. Fait remarquable mais somme toute assez logique, la communauté d'ascendance n'a même pas besoin qu'on lui ajoute «*avec modification*». La communauté d'ascendance postulée *implique* la modification. Prenons un chat et un chien. Tous deux ont une truffe et des poils.

D'où vient le partage de cette truffe et de ces poils? Trois hypothèses s'offrent à nous. (i) Ils les partagent parce que Dieu le leur a donné. Mais cela ne rentre plus dans le contrat intellectuel que les scientifiques ont collectivement passé avec les connaissances depuis la fin du XVIII<sup>e</sup> siècle (LECOINTRE, 2007), ou bien ils jugent cette réponse insuffisante<sup>1</sup>. (ii) Ils les partagent parce qu'ils vivent au même endroit, et donc les effets de l'environnement sont les mêmes sur eux. Cependant, il est facile de trouver une troisième espèce possédant truffe et poils, qui vit à des milliers de kilomètres de là et qui n'a jamais croisé ni un chat ni un chien. Un ours blanc, sur sa banquise, par exemple, habite un endroit où ne peuvent absolument pas vivre les chats. Ce n'est donc pas le seul milieu qui est directement responsable d'un tel «*façonnage*» des espèces. (iii) Enfin, troisième option, ils peuvent posséder truffe et poils parce qu'ils font des petits ensemble. Mais on sait bien d'expérience que les chats et les chiens ne font pas de petits ensemble. La réponse est transférée dans le passé : ils ont fait, *jadis*, des petits ensemble. C'est l'hypothèse de l'ascendance commune. Des ancêtres communs ont possédé truffes et poils et les ont légués au chien et au chat actuels (c'est aussi l'ancêtre de l'ours blanc, d'ailleurs). Ces ancêtres étaient-ils des chats? Si oui, les descendants se sont forcément transformés sur le trajet généalogique des chats ancestraux vers les chiens, sinon les chiens seraient chats. Ces ancêtres étaient-ils des chiens? Si oui, les descendants se sont forcément transformés sur le trajet généalogique des chiens ancestraux vers les chats, sinon les chats seraient chiens. Ces ancêtres étaient-ils autre chose? Si oui, les descendants se sont forcément transformés sur les deux trajets généalogiques de ces ancêtres vers les chiens et vers les chats, sinon les chiens et les chats ne seraient pas ce qu'ils sont. Bref, l'ascendance commune comme explication du partage d'attributs par des êtres vivants qui ne se croisent pas entre eux *implique* la transformation, quel que

1. Darwin écrivait dans le chapitre XIII de *L'Origine* : «*They believe that it [the Natural System] reveals the plan of the Creator; but unless it be specified whether order in time or space, or what else is meant by the plan of the Creator, it seems to me that nothing is thus added to our knowledge.*»

soit le processus de cette transformation. Il suffit de dire «ascendance commune», l'adjectif «commune» jouant le rôle de «avec modification».

Autre fait remarquable, l'ascendance commune dans les sciences des patrons n'a même pas besoin d'un processus particulier. Il suffit que les espèces se transforment au cours du temps, et dérivent les unes des autres, pour qu'elles soient affiliées par voie généalogique entre elles et qu'on puisse prendre ce fil généalogique pour guide classificatoire. Autrement dit, on n'a pas besoin de savoir si le processus des transformations est la sélection naturelle, la dérive génétique ou un processus lamarckien, ou un mutationnisme saltationniste, ou que sait-on encore. En fait, le processus fin de la transformation *n'est pas directement convoqué* pour concevoir une généalogie théorique des espèces ni pour en faire de la classification. Pour faire de la classification, ou pour justifier le «système naturel», la seule généalogie suffit. Un autre processus serait-il découvert pour expliquer pourquoi les générations organiques se suivent et ne se ressemblent pas, il n'y aurait pas une ligne à changer à ce paragraphe de Darwin. D'ailleurs, c'est ce que nous enseigne l'histoire des sciences. MAYR (2004) et GAYON (1992) précisent que si bien des scientifiques admirent dès le début le «*descent with modification*», peu d'entre eux adhèrent dès le début à l'hypothèse de sélection naturelle, laquelle mit soixante-dix ans pour passer de l'hypothèse très plausible au fait expérimental indiscutable. *En somme, là où l'on parle, dans les sciences des processus, de «descendance avec modification», dans les sciences des structures, l'ascendance commune (sa conséquence) suffit.*

L'ensemble de ces considérations nous amène à proposer la traduction suivante. Le premier «*descent*», puisqu'il est suivi de «*with modification*», est à traduire par «descendance avec modification» : il s'agit du cours généalogique au cours duquel se produisent les modifications. Le système naturel est légitime s'il prend en compte ce phénomène. Par contre, le second «*descent*» est clairement à traduire par «ascendance». Il n'est pas suivi de «*with modification*» (puisque l'implique déjà, s'agissant

d'espèces distinctes) et réfère explicitement aux parents communs (parents au sens des géniteurs passés). Afin de restituer la logique du texte, la traduction suivante semble s'imposer :

[...] *que le système naturel est fondé sur la descendance avec modification ; que les caractères que les naturalistes considèrent comme montrant les véritables affinités entre une ou plusieurs espèces sont ceux qui ont été hérités d'un parent commun ; ceci étant, toute vraie classification est généalogique ; que la communauté d'ascendance est le lien caché que les naturalistes ont inconsciemment recherché, et non quelque plan inconnu de création, ni l'énonciation de propositions générales, ni les seuls regroupement ou séparation d'objets plus ou moins ressemblants.* (Ma traduction<sup>2</sup>.)

Enfin, on remarquera que Barbier a ménagé ses lecteurs. Le texte de Darwin utilise une négation pour récuser l'emploi du créateur ou l'emploi d'une méthode de tri agnostique pour justifier les liens que les scientifiques seront amenés à construire pour rendre raison du partage des attributs entre espèces. Au lieu d'une négation, potentiellement problématique sur le plan de la recevabilité sociopolitique du texte, il emploie habilement «*sous prétexte de*», ce qui produit un effet de superposition. En effet, alors que dans le texte de Darwin la filiation remplace la création, dans la traduction qu'en fait Barbier, le recours à la création est un prétexte d'accès à la filiation. À ce sujet, Darwin est très clair dès la troisième page de son chapitre XIII : la justification, ou l'explication du système naturel par un plan de création doivent être remplacés par une justification et une explication par une proximité d'ascendance, parce que celle-ci est «la seule cause connue de la similarité entre les êtres vivants».

2. Ces difficultés de traduction portent encore en germe des incompréhensions entre chercheurs. Mais que dire de la même erreur de traduction, encore plus grossière, lorsque l'on considère la traduction du titre du livre de Darwin de 1871 *The Descent of Man* par *La Descendance de l'homme*, ce qui est une aberration au vu du contenu ! (voir TORT, 1996, tome 1, p. 1189-1191). Pour une traduction en accord avec l'intelligence du texte, on se reportera à la traduction coordonnée par Michel Prum et Patrick Tort parue en 1999 aux éditions Syllepse sous le titre *La Filiation de l'Homme* (DARWIN, 1999).



## INTERROGER LE GRAPHE CONNEXE NON CYCLIQUE (GCNC)

Depuis les années 1960 les systématiciens utilisent des graphes connexes non cycliques (GCNC) aux significations très diverses (voir DARLU & TASSY, 1993). Au sein même de la communauté des systématiciens, il y eut un débat autour des années 2000 pour déterminer si la théorie de l'évolution était requise pour l'exploitation des logiciels de construction d'arbre d'inspiration cladistique, c'est-à-dire ce qu'on appelle aujourd'hui la parcimonie informatisée (BROWER, 2000). Rigoureusement parlant, l'algorithme cladistique, en choisissant l'« arbre » (le GCNC) le plus parcimonieux en hypothèses de transformations des caractères le long de ses branches, maximise la contiguïté des états de caractères identiques dans l'« arbre ». Le GCNC ne fait que montrer « qui partage quoi avec qui ». Si référence à l'évolution il y a, c'est dans la tête de celui qui a construit le GCNC, alors qualifié d'arbre. *A posteriori*, l'interprétation de ce cladogramme à la lumière de la théorie de l'évolution transforme le « qui partage quoi avec qui » en « qui est plus apparenté à qui ». *A posteriori*, tout bon biologiste s'y retrouve donc. Mais c'est concernant ce qui se passe *a priori* que le débat fit rage. Certes, on peut se demander pourquoi on construirait une matrice de caractères si ce n'est pour fonder des hypothèses d'homologie primaire (de PINNA 1991) que l'on voudrait tester ; mais qui dit « homologie primaire » dit-il évolution ? En effet, les homologies primaires n'indiquent que des ressemblances associées à un principe des connections dont certains des pères, comme Richard Owen ou Georges Cuvier, étaient, faut-il le rappeler, fixistes. Ce rappel indique juste que le principe des connections, sur le plan technique, ne requiert pas le transformisme, même si aujourd'hui tous ceux qui l'utilisent inscrivent l'évolution dans *leur intention*. Le débat se trouve donc reporté sur ce que les biologistes ont dans la tête lorsqu'ils codent leur matrice de caractères. La plupart d'entre eux ne savent généralement pas pourquoi – en termes de relations de cause à effet *documentées* entre deux états de caractères présumés homologues – ces deux états de caractères se

ressemblent. Par conséquent, l'homologie primaire figurant dans la matrice repose bien sur un pari d'où l'évolution est techniquement absente (les données moléculaires et leurs modèles traités par des approches probabilistes sont laissées de côté dans cette discussion qui ne concerne que la cladistique). C'est la raison pour laquelle la plupart des matrices destinées à un traitement par parcimonie sont exemptes de pondération des caractères. Et ce n'est pas un hasard si tous les débats tournant autour de la légitimité des pondérations sont directement connectés à celui-ci. En revanche, la plupart des biologistes savent généralement pourquoi un caractère donné sera *exclu* de la matrice. La plupart des systématiciens excluront en effet de poser comme supposément homologues (homologie primaire) deux états de caractères pour lesquels on peut *présumer*, ou mieux, *documenter* une origine évolutive distincte (par l'embryologie, par l'anatomie elle-même, etc.). Ainsi, paradoxalement, c'est donc, sur le plan technique et pas seulement sur le plan des intentions, *à travers ce qu'elle ne contient pas* qu'une matrice d'homologies primaires nous parle d'évolution : elle réunit des caractères qui mettent en relation des états qui pourraient éventuellement avoir été acquis par voie d'ascendance commune.

Alors, la maximisation de la contiguïté des états identiques de caractères est-elle bonne à tout faire ? Gareth NELSON & Norman PLATNICK (1981) pensent que oui. Au sujet de la systématique phylogénétique de HENNIG (1966), Ils écrivent :

*« If so, Hennig's system would be understandable not merely as the theory of « phyletic » taxonomy but as the general theory of taxonomy of whatever sort. »*

En fait, la maximisation de la contiguïté des états de caractères identiques ne relèverait que d'un principe de cohérence maximale qui ne relève, lui-même, que de la rationalité : en sélectionnant l'arbre le plus parcimonieux, on sélectionne en même temps l'arbre le plus cohérent. En effet, plus les caractères sont cohérents entre eux dans l'arbre choisi, et moins il est besoin d'ajouter des transformations surnuméraires destinées à rendre compte des incohérences. Ce n'est pas un hasard si le nombre d'événements que requiert



un cladogramme est directement proportionnel à son « indice de cohérence » (*Consistency Index*). Cohérence et parcimonie sont une seule et même mesure. C'est au nom de l'universalité du principe de maximisation de la cohérence de nos explications que l'algorithme de parcimonie se voit conféré une portée universelle. Le principe de parcimonie est partie intégrante de l'explication rationnelle en sciences (LECOINTRE, 2007), et donc forcément mobilisé pour expliquer scientifiquement toute ressemblance. Il n'y a, alors, pas besoin d'évolution, effectivement, pour justifier l'emploi de cet algorithme. Il faut juste que nous ayons de bonnes raisons de mesurer la cohérence d'un jeu de caractères entre eux, que nous voulions utiliser un graphe connexe non cyclique pour exprimer une hiérarchie (qui revient à une série d'ensembles emboîtés les uns dans les autres – cf. HENNIG, 1966), quelle que soit la théorie explicative sous-jacente compatible avec ces deux modes de représentation d'une hiérarchie.

On devine que cette question est importante concernant l'exportation des algorithmes utilisés pour l'analyse cladistique informatisée vers d'autres champs d'investigation que la systématique : linguistique (BEN HAMED, 2009), musicologie comparative et biochimie comparative (CUNCHILLOS & LECOINTRE, 2007; TËMKIN & ELDREDGE, 2007), pour ne citer que quelques exemples. Dans ces approches comparatives d'objets qui ne sont pas des espèces (langues, instruments de musique, voies métaboliques), l'importation d'algorithmes cladistiques explicitant « qui partage quoi avec qui » n'auraient pas besoin d'être justifiées selon NELSON & PLATNICK (1981). Pour d'autres auteurs, l'utilisation de tels algorithmes requiert au minimum le fameux principe du « *descent with modification* » de Darwin. En effet, ce principe donne un sens particulier au fait de maximiser la contiguïté des états de caractères identiques. Ce principe stipule que deux états de caractères trouvés chez des espèces qui ne se croisent plus aujourd'hui doivent en première instance être interprétés comme ayant été légués par voie d'ascendance commune. Ils ne se croisent plus entre eux aujourd'hui, certes, *mais* ils doivent s'être croisés entre eux hier : il fut un temps où ils avaient un ancêtre commun. Maximiser la contiguïté de

deux (ou plus) branches portant un état de caractère donné revient à réunir ces branches de manière à leur conférer en amont un tronc commun sur lequel siège une hypothèse d'ancêtre commun. Notons que le « *descent with modification* » implique de postuler qu'il y a eu, dans le passé, une transformation des êtres (et de leurs attributs) génération après génération (ou « évolution » dans son sens le plus minimal). Cette succession de générations n'est autre qu'une *généalogie* des individus convoquée sur le plan théorique (voir section suivante). Mais la charge en postulats s'arrête là.

Nous ne nous engagerons pas ici plus avant dans le débat qui consiste à savoir si l'analyse cladistique requiert l'idée d'évolution sous sa forme « *descent with modification* », ou si cette idée est seulement souhaitable, facultative, « pas nécessaire » (BROWER, 2000), proscrite, selon que l'on se réfère au programme général de recherche de toute systématique *biologique* ou que l'on tienne la parcimonie comme seule propriété de la rationalité de l'observateur, ou encore si l'on se restreint strictement à la perception agnostique des similitudes et au fonctionnement d'un algorithme. Inscrivons pour le moment dans ce débat les différentes méthodes de construction de GCNC : pour les systématiciens de la première école (BROWER, 2000), les approches probabilistes sont chargées de présupposés superflus. Les méthodes de distances, quoique dans leur version la plus frustrée dotées de l'agnosticisme requis à l'égard du cadre théorique, sont récusées comme un traitement grossier de la similitude, récusation qui ne tient pas à l'agnosticisme théorique des approches purement phénétiques, mais à ses imperfections techniques et algorithmiques. Pour les systématiciens de la seconde école, les méthodes de distances sont récusées parce que, traduisant des degrés de similitude globale, elles sont a-théoriques (sauf pour les distances corrigées par des modèles d'évolution) et ont renoncé à l'homologie secondaire (ceci concerne en revanche tous les types de distances). Les méthodes probabilistes, elles, sont acceptées comme méthodes phylogénétiques parce qu'elles tentent de rendre explicites leurs présupposés sur l'évolution et fournissent les solutions aux problèmes d'homologie.

Les raisons pour lesquelles on utilise un GCNC sont-elles les mêmes que celles pour lesquelles il convient d'utiliser un graphe connexe cyclique (GCC)? Le GCC maximise la contiguïté des états identiques mais tracent des lignes supplémentaires pour rendre compte des ambiguïtés issues de la contradiction entre les caractères d'une même matrice. En cela, les GCC sont autant a-théoriques pour ceux qui le souhaitent que l'est pour eux le GCNC. Pour les biologistes, utiliser un GCC peut avoir une véritable signification biologique lorsque celui-ci rend compte de phénomènes de transferts d'une partie d'un organisme en un autre, ou d'hybridations. Leur utilisation est possible par exemple lorsque de nombreux transferts horizontaux sont postulés (DOOLITTLE & BAPTESTE, 2007; LOPEZ & BAPTESTE, 2009). Mais nous y reviendrons.

## INTERROGER L'ARBRE GÉNÉALOGIQUE THÉORIQUE

### De quoi parle la figure de Darwin (1859)?

Rétrospectivement, on peut dire que la seule figure de *L'Origine des espèces* est à la fois support d'explication d'un processus de diversification des espèces (dans le chapitre IV) et support d'illustration de leur filiation servant à l'explicitation d'un programme classificatoire (dans le chapitre XIII). Cette figure est un modèle qui vaut pour toutes les espèces et, transposé dans le registre classificatoire, structure l'argumentation pour tous les niveaux taxonomiques. Il y a donc, par effet d'emboîtement produit sur la même figure en une sorte de zoom arrière imaginaire, une portée modélisatrice considérable. Cette figure pose un certain nombre de conjectures, comme les expose GAYON (2009) :

- les espèces se *modifient* au cours du temps ;
  - cette modification est *graduelle* ;
  - beaucoup d'espèces s'*éteignent* ;
  - généralement, les espèces qui ne s'éteignent pas se *scindent* en plusieurs ;
  - une fois scindées, elles *divergent* indéfiniment et graduellement à leur tour ;
- on peut remplacer le mot « espèce » ci-dessus par des entités taxonomiques de plus petit rang ou bien de plus haut rang, si bien que la classification sera entièrement déterminée par la généalogie. Par conséquent, il n'y a qu'une seule grande généalogie théorique du vivant, étendant ses divergences et ses extinctions le long d'un temps relatif symbolisé par des barres horizontales, qui seront tantôt des tranches de temps mesuré en milliers de générations, tantôt des couches géologiques ;
  - les catégories taxonomiques (les rangs) sont arbitraires mais leur assignation doit suivre des règles (chapitre XIII, voir ci-dessous) ;
  - le processus entier de diversification biologique se réduit au processus de différenciation au niveau élémentaire des variétés et des espèces ; dont les taxons de rang supérieur ne sont que des boîtes fabriquées par nous, fondées sur le partage des attributs, lequel n'est que le résultat du processus à petite échelle.

Cette figure ne se lit donc pas comme une phylogénie actuelle, puisqu'elle traduit un processus théorique dans lequel figurent des relations d'ancêtres à descendants, mais davantage comme le *paramétrage*, exposé par dessein, de la forme que doit détenir une généalogie théorique du vivant, de laquelle on peut, secondairement et bien entendu rétrospectivement, dresser un arbre phylogénétique moderne qui ne retranscrirait que des degrés relatifs d'apparementent. Comme le remarque GAYON (2009) :

*D'un point de vue méthodologique, le diagramme de Darwin n'est certainement pas un mode de représentation neutre des faits, comme l'a trop suggéré dans le passé l'expression de « fait général de l'évolution ». C'est une authentique construction théorique. [...] Ce diagramme a l'allure d'une généralisation descriptive, mais ce n'est pas le cas. C'est plutôt un pari heuristique sur la forme et l'allure générale des phénomènes qui manifestent l'évolution by and large et requièrent des explications causales.*

Comme tout modèle en sciences, cet « arbre » de Darwin a été discuté durant les cent cinquante ans qui ont suivi sa publication. Pour être plus précis,

certaines des conjectures ci-dessus ont fait l'objet de contestations que nous résumerons avec GAYON (2009) :

- l'évolution n'est pas nécessairement graduelle ;
- les espèces ne sont pas toujours irréversiblement divergentes après leur scission (surtout chez les micro-organismes, DOOLITTLE & BAPTESTE, 2007 ; LOPEZ & BAPTESTE, 2009, voir section suivante) ;
- on a contesté l'applicabilité du diagramme à tous les niveaux taxonomiques ;
- on a voulu plaider pour des mécanismes macro-évolutifs qui seraient distincts des processus micro-évolutifs.

Fait remarquable, comme les lignées sont organisées le long du temps, les êtres vivants actuels sont tous au même niveau. Cette figure est véritablement un modèle scientifique, et en cela n'incorpore pas de valeurs. Darwin avait, par cette figure, véritablement rompu avec le scalisme, lequel reviendra « par la fenêtre » avec les grades, lesquels ne seront évacués par la suite qu'avec la généralisation de la systématique phylogénétique de Willi HENNIG (1950, 1966) et ses descendants informatisés. Ce n'est pas Darwin qui dessinera le premier arbre phylogénétique pourvu d'un tronc central et d'un sommet auquel culminera l'homme, mais HAECKEL (1874), dans un abus métaphorique. Par rapport à son contenu comme à sa rigueur, les arbres de HAECKEL (1866, 1874) sont bien loin de l'arbre de DARWIN (1859).

### La forme de l'arbre de la vie

Depuis que des transferts « horizontaux » de matériel génétique ont été prouvés entre lignages bactériens en principe généalogiquement séparés depuis plus de deux milliards d'années, entre bactéries et champignons, entre bactéries et plantes érigées, l'image de généalogie théorique que publia Darwin dans *L'Origine des espèces* a été corrigée. La métaphore généalogique reste de mise, mais la forme de l'arbre théorique de la vie s'est complexifiée. En effet, les lignages ne sont plus considérés comme nécessairement divergents, des réticulations apparaissent : des échanges partiels horizontaux s'élaborent entre

branches parfois séparées depuis fort longtemps, ou de nouvelles espèces apparaissent par hybridations d'espèces mères (par exemple, le cas est bien documenté comme apparaissant de manière récurrente en milieu naturel chez les tournesols). L'arbre généalogique de la vie, le fameux « *tree of life* », généalogie théorique divergente (GTD) tendrait alors à devenir une généalogie théorique réticulée (GTR), au moins dans certaines de ses parties (DOOLITTLE & BAPTESTE, 2007 ; LOPEZ & BAPTESTE, 2009). Pour certains organismes unicellulaires, la réticulation, ou le lignage horizontal, dominerait le lignage vertical. Le « *descent with modification* » n'est plus que composite et il a, en quelque sorte, changé de direction. Les procédés d'acquisition changent, la modification subséquente reste possible. Il résulte de ce schéma théorique que, compte tenu du fait que ce sont le plus souvent des portions de génomes qui sont échangés horizontalement, l'histoire généalogique des gènes peut très bien ne pas être l'histoire généalogique des espèces qui les portent. Le problème se pose avec d'autant plus d'acuité que les reconstructions phylogénétiques impliquent de reconstituer des divergences anciennes. Il convient de ne pas confondre le schéma généalogique théorique (l'ontologie) et la phylogénie (la méthodologie). Rappelons qu'un arbre phylogénétique, en tant qu'outil, ne dit en première instance que « qui partage quoi avec qui ». C'est le fond théorique qui fournit une première couche interprétative en permettant de passer de l'interprétation précédente à « qui est apparenté à qui ». Une troisième couche interprétative prend en compte la possibilité de transferts horizontaux. Si les transferts horizontaux sont susceptibles de s'être produits entre des espèces dont on cherche les relations d'apparement, il faut mener l'investigation sur plusieurs gènes du génome et produire pour ces mêmes espèces leurs phylogénies à partir de chaque gène, séparément. Il faut obtenir pour ces espèces autant de phylogénies que de gènes. L'arbre phylogénétique garde sa puissance heuristique, mais il faudra lire ces arbres comme des arbres de gènes et non plus en première instance comme des arbres d'espèces. C'est de la confrontation des différents arbres de gènes entre eux, et de leurs discordances

éventuelles, que seront interprétés certains des apparentements comme issus de transferts et certains autres comme issus d'ancêtres communs aux espèces (cf. par exemple ESCOBAR-PARAMO *et al.*, 2004). Le fait d'avoir changé l'image généalogique théorique ne remet pas en cause la puissance heuristique du graphe connexe non cyclique. L'utilité de graphes connexes cycliques, venus récemment en complément, est actuellement discutée (HUSON & BRYANT, 2006). Cependant, pour des organismes unicellulaires dont seulement 30 % du génome a des chances d'avoir été acquis par voie verticale, l'intensité des transferts horizontaux a été telle que, au-delà de la reconstitution d'une histoire de gènes, on en finit par ne plus savoir de quoi, au-delà des gènes, on fait l'histoire. Ceci est vrai que l'on utilise des graphes connexes cycliques ou des graphes connexes non cycliques. L'identité de l'organisme mosaïque fait question, en même temps que le cahier des charges de sa classification. En effet, s'il ne subsiste quasiment plus rien d'acquis verticalement qui soit identifiable, on s'oriente vers une classification de structures acquises horizontalement dont la présence nous parle davantage des milieux par lesquels les ancêtres sont passés ou subsistent. La similitude globale, si elle est utilisée, ne vous raconte plus les degrés d'apparentement mais une succession de milieux à mémoire courte. Ce n'est pas la méthode phylogénétique qui est en cause mais son interprétation dans un cadre où se perd la trace phylogénétique.

### **Courroie de transmission entre processus et patrons**

Le modèle de Darwin présente donc les paramètres constitutifs de la forme d'une grande généalogie théorique du vivant. Ce modèle fonctionne comme la courroie de transmission entre la science des patrons et la science des processus. Cette division réfère au type de questions que pose le chercheur. Les sciences des patrons cherchent à structurer des agencements entre entités du vivant, de manière à les rendre intelligibles à l'aide de concepts et de mots de portée générale. Les sciences des processus mettent en évidence des relations de cause à effet. Les deux types de sciences utilisent des modalités de la preuve

qui sont distinctes (LECOINTRE, 2009) sans s'ignorer mutuellement pour autant. Il est important de noter que cette distinction vaut pour l'opérationnalité de la recherche, mais n'a plus nécessairement lieu d'être dans une phase de synthèse des connaissances. Le chercheur qui travaille sous l'un des deux régimes doit en être conscient, et cette conscience tient à la nature de l'interrogation qui l'anime (*i.e.* la question qui sera posée dans son article de recherche) et au type de preuve qui sera mis en œuvre pour y répondre. Les prémisses et les données sur lesquelles le chercheur se fonde, elles, peuvent se nourrir indistinctement dans les deux champs. En résumé, le point de départ du chercheur relève indistinctement des deux champs, la question posée, la mise en preuve et la réponse sont spécifiques, la phase de synthèse des connaissances est indistincte. Pour revenir au modèle généalogique darwinien, il fonctionne comme une courroie de transmission entre processus et patrons parce que :

- il indique à quoi il faut s'attendre, comme forme généalogique, si les espèces se transforment ; le seul « bagage » minimal en termes de processus étant l'héritabilité de ce qui varie, la scission longitudinale du flux généalogique et surtout, dès lors, le principe de divergence entre lignées désormais séparées (la sélection naturelle ou toute autre forme de stabilisation d'un trait dans une espèce n'ont pas réellement besoin d'être convoqués) ;
- il indique ce à quoi il faut s'attendre comme répartition d'attributs dans la nature actuelle, suite à ce déploiement généalogique, et donc quels paris sont légitimes : si deux espèces ne se croisant pas aujourd'hui possèdent des attributs en commun, c'est qu'elles les auront acquis par voie d'ascendance commune ;
- il fournit le guide, le squelette qui charpentera le « cahier des charges » de la classification. En effet, cet arbre est généalogique et théorique, mais c'est dans le texte du chapitre XIII qu'on trouve la recommandation de monophylie qui sous-tendra le programme hennigien, comme nous allons le voir dans la section suivante.

### La contrainte de monophylie est-elle incluse chez Darwin (1859) ? : autres difficultés de traduction

Celui qui fixera réellement – c'est-à-dire explicitement et durablement – le cahier des charges des classifications sera Darwin (cf. citation, section « Difficultés de traduction »). Ce texte est remarquable à plus d'un titre. Tout d'abord, il fonde un programme unique à la classification : la proximité généalogique. Ensuite, il réalise un changement de paradigme : la création n'est plus convoquée et dès lors, l'application de ce programme doit *en principe* donner une place à l'homme dans la classification en vertu de ses affinités généalogiques, et non comme créature parfaite culminant au sommet d'une création aux intentions obscures. Enfin, ce texte contient en germe une autre révolution qui ne portera ses fruits qu'au XX<sup>e</sup> siècle. En effet, les logiques divisives ou agglomératives fondées sur la ressemblance globale sont également récusées (« *the mere putting together and separating objects more or less alike* »). Ce n'est pas la ressemblance globale qui doit fonder la classification, mais la prise en compte de ceux des caractères qui sont hérités d'un parent commun. La logique d'une approche par caractères est déjà, en quelque sorte, préconisée (même si elle ne deviendra opérationnelle que dans les années 1950) en opposition à une approche globalisante dont on sait aujourd'hui qu'elle mélange les caractères récemment hérités, les caractères trop anciens, mais surtout les caractères convergents ou les réversions qui conduisent à des ressemblances non héritées d'un parent commun (ou homoplasies). Curieusement, c'est sans tenir compte de ce texte qu'un programme de classification par ressemblance globale sera importé en biologie par l'école phénétique des années 1958-1970. Mais dans la même période, l'école phylogénétique de HENNIG (1950) réalisera pleinement ce que ce texte implique.

Les groupes, ou taxons, ont donc un rapport direct à des lignées phylétiques. Cependant, en 1859, DARWIN ne disposait pas du mot « phylogénie ». Ce mot fut forgé par Ernst HAECKEL en 1866. Darwin l'utilisa ensuite dans ses cinquième et sixième éditions pour parler de généalogie. Mais il était clair pour

Darwin qu'il n'existait pas, pour les êtres vivants, de registres d'état-civil permettant d'identifier les ancêtres individuellement, et que cette généalogie – forcément incomplète – ne pouvait être que partiellement reconstituée par un exercice de comparaison des êtres actuels. Pour désigner cet exercice, nous parlons aujourd'hui de « reconstruction phylogénétique » et, si un programme classificatoire s'en suit, de classification phylogénétique. Darwin utilisait donc le mot « généalogie » sur le plan théorique comme nous le faisons, sur lequel il a plaqué tardivement le terme de Haeckel « phylogénie » sans apporter de réponses quant au moyen d'accéder à cette phylogénie. Nous utilisons aujourd'hui le mot « phylogénie » comme cette partie de généalogie que nous sommes capables de reconstituer partiellement à partir de descendants concrets, comblant une pratique heuristique qui manquait à Darwin. Malgré les souhaits darwiniens en matière de classification, entre 1859, date de parution de *L'Origine des espèces*, et 1950, date de parution de la *Systématique phylogénétique* de l'entomologiste allemand Willi Hennig, on a mélangé les relations d'ancêtres à descendants et des relations de parenté (parenté au sens de degré de cousinage relatif : qui est plus proche de qui ?) en même temps que dans les classifications persistaient des groupes qui ne reflétaient pas seulement la phylogénie. Ces groupes, les grades, reflétaient un niveau de développement de la complexité des organismes, séparés entre eux par des « sauts adaptatifs » qui gommaient les relations d'apparement.

Par exemple, parmi les cinq classes de vertébrés de la zoologie traditionnelle, trois d'entre elles sont des grades. Les poissons sont les vertébrés qui ne sont pas « sortis des eaux » (ils n'ont pas le membre chiridien) ; les amphibiens sont des tétrapodes non émancipés du milieu aquatique (ils n'ont pas l'œuf amniotique) ; les reptiles sont les amniotes sans poils ni plumes. Comme les grades sont des degrés de complexité, ils sont fondés sur des absences d'attributs. Comme ils ne tiennent pas compte de la phylogénie, certains de leurs membres sont plus apparentés à des organismes extérieurs au grade qu'à leurs collatéraux dans le grade. Hennig dira qu'ils sont paraphylétiques. La truite est plus apparentée

à nous qu'à un requin. Les amphibiens seymouriamorphes sont plus apparentés aux amniotes qu'aux grenouilles. Les crocodiles sont plus apparentés aux oiseaux qu'aux lézards et les gorgonopsiens sont plus apparentés aux mammifères qu'aux tortues. Si les chercheurs ont continué à fabriquer des taxons qui ne tiennent pas compte des degrés d'apparementement après les recommandations de DARWIN (1859), il faut se demander si Darwin avait été aussi clair qu'on aurait pu le souhaiter. En effet, les recommandations que fait Darwin dans son chapitre XIII sur la classification, aidé de la figure centrale du livre, montrent qu'on doit classer tous les descendants d'un même ancêtre dans un groupe. Darwin recommande donc de faire ce que HENNIG (1950) appellera plus tard des groupes monophylétiques. Cependant, Darwin perçoit la difficulté liée à des groupes qui accumulent des spécialisations sur leur lignée propre, spécialisations qui, si elles sont nombreuses, pourraient occulter les signes de l'apparementement. Faut-il faire alors un groupe spécial pour cette lignée, et la couper ainsi de son ascendance ? Ou bien ne pas tenir compte de cette spécialisation et mettre l'accent sur les caractères qui relie la lignée aux autres lignées ? La question n'est pas anodine, car elle sous-tend les pratiques classificatoires du siècle qui suivra Darwin, et sera au cœur du changement de paradigme qui aura lieu dans les années 1970 au moment où, HENNIG (1950) ayant été traduit en anglais (HENNIG, 1966), les systématiciens commencent à changer leurs méthodes de travail. Prenons un exemple. Devant constituer de grandes classes de vertébrés, et face à ce qui nous semble une énorme accumulation de spécialisations que requiert l'aptitude au vol battu chez les oiseaux (plumes asymétriques, fourchette, bréchet, os creux, sacs aériens, etc.), devons-nous extirper les oiseaux de leur ascendance reptilienne indéniable en les élevant au rang de classe, et en constituant donc en parallèle un marchepied qu'on appellera classe des « reptiles » (le grade des reptiles étant donc constitué d'amniotes qui ne volent pas) coupé des oiseaux, ou au contraire assignerons-nous un rang inférieur aux oiseaux afin de les laisser au sein des « reptiles », ainsi maintenus dans leur intégrité phylogénétique ? On doit se rendre à l'évidence avec NELSON (1972), Darwin préconisait

la monophylie en favorisant la seconde option. Dans son chapitre XIII des première et seconde éditions de *L'Origine*, Darwin est très clair :

« S'il pouvait être prouvé que l'Hottentot descend des Noirs, je pense qu'il serait classé dans le groupe des Noirs, quelle que soit la quantité par laquelle il diffère des Noirs en termes de couleur ou autres caractères importants<sup>3</sup>. »

Ici c'est clairement la monophylie qui est préférée (Darwin reprend le même raisonnement à propos de pigeons dans la même page, figurant dans la traduction de Barbier). Reprenons le raisonnement (et la même phrase) pour les classes de vertébrés : « S'il pouvait être prouvé que le poulet descend des reptiles, je pense qu'il serait classé dans le groupe des reptiles, quelle que soit la quantité par laquelle il diffère des reptiles en termes de couleur ou autres caractères importants ». Ce texte consiste alors à faire des reptiles un groupe monophylétique en y incluant les oiseaux, ce que préconisera la systématique phylogénétique après 1966. Aujourd'hui encore, si l'on souhaite maintenir les reptiles dans la classification, il faut y inclure les oiseaux (LAURIN, 2008 : 18). C'est le même raisonnement que nous tenons lorsque nous considérons les oiseaux comme des dinosaures actuels.

Darwin, dans son chapitre XIII, ne préconise donc pas les grades : pour classer, priorité doit être donnée à la filiation, à l'ascendance commune, et pas aux degrés de spécialisations ultérieurs propres à des lignées uniques, aussi spectaculaires soient-elles. Pourquoi n'a-t-on pas fait des groupes monophylétiques tout de suite ? TASSY (1991 : 48) assigne ce retard à une ambiguïté de Darwin lui-même :

*Les systématiciens prédarwinien, Darwin lui-même et nombre de ses successeurs ont utilisé la somme des modifications comme un outil taxinomique de grande efficacité. Dans l'Origine, Darwin précise : « Je crois que l'arrangement des groupes dans chaque classe, d'après leurs relations et leur degré de subordination mutuelle, doit, pour être naturel, être rigoureusement*

3. Ma traduction, ce passage ne figurant pas dans la traduction française de Barbier, qui est une traduction parue en 1876 de la sixième édition de *L'Origine des espèces*.

généalogique ; mais que la somme des différences dans les diverses branches ou groupes, alliés d'ailleurs au même degré de consanguinité avec leur ancêtre commun, peut différer beaucoup, car elle dépend des divers degrés de modification qu'ils ont subis ; or, c'est là ce qu'exprime le classement des formes en genres, familles, sections ou ordres. » Et Darwin conclut : « Le système naturel ramifié ressemble à un arbre généalogique ; mais la somme des modifications éprouvées par les différents groupes doit exprimer leur arrangement en ce qu'on appelle genres, sous-familles, familles, sections, ordres et classes. » La somme des modifications n'est pas un concept généalogique, elle ne nous donne pas la filiation ; elle est donc typologique. Mais, semble-t-il, Darwin admet qu'elle puisse être à la source de l'assignation des groupes à telle ou telle catégorie de la classification. De l'inclusion d'un groupe à tel niveau de la hiérarchie à l'identification même du groupe, il n'y a qu'un pas. Dans sa conclusion, Darwin use même du vocable « arrangement » pour l'assignation des groupes aux différentes catégories – alors qu'il aurait dû utiliser celui de « ranking », en français : « catégorisation ». Cette ambiguïté formelle et cette concession à une pratique typologique de la classification seront lourdes de conséquences. On peut y voir la cause première de ce qui sera considéré par le néodarwinisme du *xx<sup>e</sup>* siècle comme le meilleur mode de représentation de la phylogénie : des groupes ancestraux dont les membres sont peu divergents entre eux, et qui donnent naissance à des groupes qui éprouvent divers degrés de modification, c'est-à-dire de divergence.

En fait, Darwin ne fut pas aussi ambigu que Tassy le dit, et il se trouve que Darwin *utilisa précisément* le mot « ranking » que Tassy appelle de ses souhaits. Cependant, c'est la traduction d'Edmond Barbier, que Tassy cite, qui est incapable de retranscrire proprement la pensée de Darwin sur ce point. Car voici ce que DARWIN écrit réellement dans les première (1859) et seconde (1860) éditions de *L'Origine des espèces* :

*I believe that the arrangement of the groups within each class, in due subordination and relation to the other groups, must be strictly genealogical in order to be natural ; but that the amount of difference in the several branches or groups, though allied in the same degree in blood to their common progenitor, may differ greatly, being due to the different degrees of modification which they have undergone ; and this is expressed by the*

*forms being ranked under different genera, families, sections, or orders.*

*The natural system in genealogical in its arrangement, like a pedigree ; but the degrees of modifications which the different groups have undergone, have to be expressed by ranking them under different so-called genera, sub-families, families, sections, orders, and classes.*

On voit clairement que la traduction de Barbier ne comprend pas le texte. Pire, il y a des inversions de rapports sur le statut de cette « somme de modifications » propre aux lignées : pour Darwin, elle doit être exprimée par quelque chose, pour Barbier elle exprime quelque chose. Pour Darwin, la somme des modifications doit être exprimée par la catégorisation, pour Barbier elle exprime leur « arrangement ». Entre fabriquer un groupe (classer, faire des « arrangements »), assigner une espèce ou un groupe d'espèces dans un groupe plus grand déjà constitué d'une part, ou assigner un rang formel (ou catégorie : famille, ordre, classe, etc.) à un groupe déjà constitué d'autre part, nous avons là des opérations distinctes. Barbier traduit « ranking » par « classement » dans la première phrase, puis par « arrangement » dans la seconde phrase, avec en plus une inversion de rôle des degrés de modifications qui passent de la voie passive à la voie active. On ne peut faire plus confus, lorsque l'on sait que « ranking » devait se traduire par « assignation d'un rang à un groupe » ou « assignation d'une catégorie à un groupe » ou, comme Tassy le remarque judicieusement, « catégorisation ». Pour la fabrication des groupes (ou l'« arrangement », ou l'action de classer), Darwin préconisait la stricte filiation, quel que soit le degré ou la somme des modifications propres (voir le passage sur les pigeons ou sur l'Hottentot). Cette somme de modifications ultérieures propres à un lignage devait être gérée, pour Darwin, en termes d'assignation de rangs (*ranking*), pas en termes de constitution de groupes (*arrangement*) ; Darwin est clair sur cette question, dans le même texte lorsqu'il traite de ce qu'il faut faire de la lignée F de sa figure au regard de ce qui est fait pour la lignée A et I. Pour rendre le propos plus accessible, prenons un exemple concret, celui des mammifères. Imaginons que le déploiement

de la lignée A de la figure de *L'Origine des espèces* soient les mammifères thériens (placentaires et marsupiaux), et le déploiement de la lignée F soit l'ornithorynque. Ce dernier possède avec la lignée A la glande lactéale à l'origine du lait, les poils et la mandibule constituée d'un seul os, l'os dentaire. Ces trois traits sont des traits qui rattachent sans ambiguïté l'ornithorynque au lignage de A, dans un groupe appelé mammifères. L'arrangement consiste à faire un seul groupe, celui des Mammalia, comprenant A et F. Cependant, l'ornithorynque est vraiment un mammifère « spécial », avec sa structure osseuse temporale, ses dents spéciales, son bec corné, ses éperons venimeux aux pattes postérieures des mâles, ses pattes palmées, sa queue plate... Comment gérer cette « somme » des modifications qu'a subi la lignée propre aux ornithorynques ? Selon la traduction de Barbier, s'il s'agit de gérer par « arrangement » : alors nous aurions fait un groupe à part des ornithorynques par une coupure divisive, afin de souligner leur exceptionnelle dérivation, et un groupe des mammifères non ornithorynques (regroupant les échidnés, les marsupiaux et les placentaires ; l'analogie est ici claire avec les poissons, qui sont des vertébrés non tétrapodes, et les reptiles qui sont des amniotes non oiseaux). Selon Darwin lui-même, cette situation est à gérer en assignation de rang. L'ornithorynque, au lieu d'avoir sa « boîte » portant l'étiquette de famille, verra assignée à cette boîte le rang de sous-classe. La traduction de Barbier donne aux spécialisations des lignées le critère de fabrication des boîtes (ce que feront les systématiciens après Darwin), tandis que Darwin souhaitait ne gérer l'empilement des spécialisations propres aux lignées qu'à l'aide de rangs assignés à des boîtes qui, elles, devaient rester régies par l'apparement. Une fois de plus la traduction de Barbier introduisit donc une grande confusion. On remarquera pourtant qu'il n'y a pas de contradiction entre les deux recommandations de Darwin. Ce n'est pas parce qu'il faut gérer les spécialisations des lignées en assignation de rang que cela conduit nécessairement à suivre une logique divisive ou à faire des grades. Darwin ne dit simplement pas comment faire, il reste au seuil d'une vraie méthodologie, et comme le souligne DUPUIS (1988 : 90-91), il laisse

ses successeurs se débrouiller au cas par cas. La vraie raison pour laquelle Darwin ne fut pas suivi dans le siècle qui succéda à la parution de *L'Origine* ne fut donc pas tant son ambiguïté quant à l'expression de son programme, mais surtout qu'il ne donna pas le mode d'emploi pour le réaliser. Il considérait presque la classification strictement généalogique comme un idéal hors d'atteinte ; attitude actant tacitement de l'absence de méthode de construction d'arbre partant d'êtres concrets. Sans construction formelle d'arbre, pas de distinction entre des ancêtres abstraits et des ancêtres concrets ; pas de solution à la tension ressentie chez Darwin entre l'arrangement « généalogique » et l'assignation de rangs (la catégorisation nomenclatoriale [ou nomenclaturale], cf. DUPUIS 1988). Même si cette attention portée à une traduction peut sembler désuète ou superflue au lecteur, TASSY (1991) a raison de signaler avec DUPUIS (1986, 1988) que l'enjeu de ce texte est considérable pour comprendre ce que les systématiciens vont faire dans le siècle suivant. L'absence de méthode permet, selon le mot de DUPUIS (1986), la survie de « procédures insuffisamment épurées » et le mélange ultérieur en systématique de taxonomies divisives et agglomératives. Mais moins techniquement, sans méthode nouvelle, et avec des traductions confuses (pour ce qui concerne le contexte français), c'est la tradition qui, par inertie, se maintint.

Existe-t-il d'autres raisons pour lesquelles les classifications biologiques ne devinrent pas purement phylogénétiques après Darwin ? Il y eut deux raisons principales.

La première tient au contexte social dans lequel Darwin émit ses idées et au rôle des fossiles. Sommés par leurs détracteurs dès 1860 – dont la plupart émettaient leurs critiques du dehors des sciences – de fournir des preuves de l'évolution biologique, les darwiniens eurent tendance à présenter les fossiles comme des ancêtres identifiés, de véritables ancêtres incarnés, preuves matérielles du déroulement évolutif, ce qui acheva la confusion entre généalogie (relations d'ancêtres à descendants : qui descend de qui, non directement accessibles à l'investigation) et la phylogénie (relations de parenté : qui est plus



proche de qui, accessibles à l'investigation). Ce fut le cas par exemple pour *Archaeopteryx*, oiseau fossile découvert en 1861, conçu comme l'ancêtre des oiseaux. Il faut comprendre que dès lors, les relations d'ancêtres à descendants furent conçues entre taxons, comme ce fut le cas dès 1866 dans les arbres des paléontologues transformistes français Albert Gaudry (sur les hyènes fossiles, cf. TASSY 1991 : 60) et allemand Franz Hilgendorf (sur les planorbes fossiles) où les fossiles attestent les filiations de par leur qualité même de fossiles. Lorsque les relations d'ancêtres à descendants seront dessinées entre taxons de plus haut rang encore, les fossiles continueront d'incarner des ancêtres concrets, constituant alors des ponts entre les grades, ou entre les grades et les clades. Le fossile comme ancêtre concret participe de la logique gradiste. Tant et si bien que dans toute l'iconographie de ces auteurs et de leurs héritiers du XX<sup>e</sup> siècle (tels A. Romer), les fossiles sont représentés aux nœuds des branches des arbres phylogénétiques, erreur que font encore aujourd'hui nombre de paléontologues et de médias.

La deuxième raison tient au fait qu'on a voulu injecter dans la classification biologique beaucoup plus que la seule phylogénie. En effet, la systématique de l'époque, que l'on qualifiera plus tard de «systématique éclectique», s'employait à traduire non seulement les «affinités évolutives», mais aussi les «sauts adaptatifs» et le «degré de complexité» réalisés par des taxons entiers appelés «grades». Par exemple, on sait depuis la fin du XIX<sup>e</sup> siècle que les oiseaux ont certaines affinités avec les dinosaures théropodes. Autrement dit, certains dinosaures théropodes (comme le *Velociraptor*) sont plus apparentés aux oiseaux qu'à n'importe quel autre reptile ; c'est-à-dire qu'ils partagent avec les oiseaux des ancêtres communs exclusifs. Cela signifie que la classe des reptiles est hétérogène sur le plan phylogénétique : elle comprend des éléments (certains dinosaures) plus apparentés à des organismes non reptiles (les oiseaux) qu'aux autres reptiles. Malgré cela, au lieu d'inclure les oiseaux dans les reptiles afin de restituer à ces derniers une partie de leur histoire, on a maintenu durant un siècle une classe des oiseaux séparée de la classe des reptiles pour souligner une

grande différence anatomique *globale* entre les deux classes. Les oiseaux ayant appris à voler, ils ont acquis sur leur propre branche de l'arbre évolutif une «somme» de spécialisations (plumes, bréchet, anatomie du membre antérieur, etc.) qui marque un «saut adaptatif», ou encore réalisent un bond en terme de «degré de complexité» (bien que cette complexité n'ait jamais été objectivement définie). Le «grade» de reptile, comme tous les grades, est fondé sur la ressemblance globale, laquelle exprime un «degré général de complexité» et se délimite par un «saut adaptatif» que réalise le groupe auquel le grade donne naissance. En soulignant ce saut adaptatif, la systématique éclectique masquait le véritable groupe-frère des oiseaux et ceci tout en sachant très bien quel était ce groupe-frère.

À titre d'information, les oiseaux, les «dinosaures», les ptérosaures, les crocodiles et un certain nombre de groupes fossiles moins populaires constituent l'ensemble des archosaures. Dans la faune actuelle, les seuls archosaures restants sont les crocodiles et les oiseaux. On notera que les reptiles n'ont pas plus d'homogénéité lorsqu'on les considère uniquement à partir de la faune actuelle : les crocodiles sont plus apparentés aux oiseaux qu'à n'importe quel groupe de reptiles actuels (tortues, lézards, serpents, rhynchocéphales). Les crocodiles partagent notamment avec les oiseaux la mandibule osseuse fenêtrée et le gésier.

Les reptiles n'existent donc que pour souligner le saut adaptatif des oiseaux, ce qu'admettait l'école gradiste. La distinction négative reptile / oiseaux masque l'apparemment exclusif crocodiles / oiseaux. *En somme, le saut adaptatif et la discontinuité de complexité sont des concepts évolutionnistes qui masquent dans la classification les véritables liens de parenté, et qui ont permis la perpétuation d'anciens groupes linnéens sous un vernis évolutionniste. Mais il y a pire, au moins à deux titres.*

Premièrement, les grades donnent une vie évolutive à des taxons de haut rang, oubliant ainsi que les taxons sont créés par le besoin du classificateur, qu'ils ne sont là que pour remplir un cahier des charges que celui-ci leur a fixé, et que, par conséquent, ils

n'ont pas de *dynamique* biologique ou évolutive. Dans la nature il n'y a que des individus, lesquels portent des attributs. Les caractères et les classifications qui en découlent sont des constructions scientifiques. La meilleure preuve de cela est simple : les classifications changent, les organismes réels qu'elles contiennent, eux, restent. Ce n'est pas parce que le concept de «poisson» n'est plus valable en systématique que les truites ou les perches ont disparu. Ce dont nous sommes capables de reconstituer l'évolution, ce ne sont pas les taxons, ni même les espèces ; le phylogénéticien ne peut reconstituer l'évolution que d'attributs. Le grade est essentialiste parce qu'il assigne à un concept classificatoire des mécanismes biologiques ; considère qu'un groupe taxonomique évolue et donne naissance à un autre, parce que l'essence du reptile se maintient dans le taxon jusqu'à un certain degré malgré l'évolution intrinsèque de ce qu'il contient ; il revient *en arrière* du nominalisme de Darwin.

Deuxièmement, les grades sont de véritables marchepieds, les nouveaux barreaux d'une échelle des êtres larvée, parce que devenue évolutionniste. Les reptiles n'existent que parce que certains d'entre eux donnèrent un jour naissance aux oiseaux, les poissons parce qu'ils donnèrent un jour naissance aux tétrapodes, les invertébrés aux vertébrés, les pongidés aux hominidés. Ainsi, bon nombre de ces groupes de la classification linnéenne devenue évolutionniste, conservés entre Darwin et Hennig, ne sont pas définis pour eux-mêmes par un attribut qui leur soit exclusif, mais par rapport à leur «devenir évolutif». Certains groupes furent même créés pour cela, portant dans leur étymologie l'idée de «gestation évolutive» : ainsi en va-t-il de tous les «pro»-quelque chose, tels les procaryotes annonçant les eucaryotes, les prosimiens annonçant les simiens, les protacanthoptérygiens annonçant les acanthoptérygiens. En fait, fonder un taxon sur son devenir est une grave faute logique en sciences de l'évolution, parce qu'aucun devenir n'est inscrit : les organismes vivants ne sont porteurs que de leur passé. Ces relations d'ancêtres à descendants entre groupes de haut rang permirent la résurgence de l'image de l'échelle des êtres, mais cette fois-ci évolutionniste, renforcée par l'intrusion

de valeurs dans le discours scientifique, par exemple le «progrès». Certains barreaux donnent naissance à d'autres. La truite (barreau des poissons) est conçue par nos journalistes comme une étape sur le chemin menant du requin à l'homme, qui est en haut. La vulgarisation des résultats de la systématique, pas toujours bien renseignée, fourmille d'expressions où la biodiversité est conçue comme rangée le long d'un seul vecteur, «de la bactérie à l'éléphant» ou «de la bactérie à l'homme», c'est selon. Cette échelle des êtres est incompatible avec l'arborescence qui réserve toute possibilité à tout organisme d'accumuler sa complexité unique sur sa branche propre, tout en préservant la possibilité de trouver les attributs signes de l'apparement sur des troncs communs. La truite possède des attributs différents de ceux de l'homme. En redescendant suffisamment bas dans l'arbre de la vie, on peut trouver à – 420 millions d'années des attributs communs aux deux qu'ont légué des ancêtres communs hypothétiques : les os dermiques tels que le pariétal, le maxillaire, le dentaire, par exemple. Autrement dit, en rompant avec l'échelle des êtres, ce n'est pas parce que la truite évolue dans une direction différente de celle de l'homme qu'on ne peut pas retrouver ce qui la lie à l'homme. Ceci est vrai pour tout le vivant. Quand on dit que l'ornithorynque ou le coelacanthé sont «primitifs», on ne raisonne pas dans un arbre, mais sur une échelle des êtres, voire en essentialiste. Dans le cadre mammalien, l'ornithorynque et autres monotrèmes sont primitifs au regard de quelques attributs, comme l'œuf à coquille dure et l'absence de tétos. Mais il fourmille d'innovations que l'on ne trouve chez nul autre mammifère : le bec corné «de canard», des dents très spéciales, des éperons venimeux aux pattes postérieures des mâles... Il n'y a pas d'espèces ou de groupes entiers qui soient primitifs. Seuls des attributs peuvent l'être. Et encore..., être «primitif» dans l'absolu ne signifie rien. Les attributs sont primitifs ou dérivés en rapport à un cadre d'échantillonnage. Pondre des œufs à coquille dure est un trait primitif dans un échantillon limité à des mammifères, mais un trait dérivé dans un échantillon de tétrapodes. Bref, c'est une valeur relative qui ne peut qualifier qu'un attribut. La plupart des discours journalistiques sur la

nature sont en rupture avec ces concepts, par la force de l'anthropocentrisme, de l'essentialisme et l'ignorance des procédures classificatoires modernes.

Le succès des grades tint donc à la persistance des groupes traditionnels qu'ils permettaient, soutenue par une lecture scaliste de la nature, puisque les grades s'empilaient les uns sur les autres, chacun soulignant, annonçant presque l'étape suivante. Cette échelle était flatteuse de surcroît puisque, doublée d'un progrès évolutif, elle consistait à mettre l'homme en haut de l'arbre de HAECKEL (1874) et s'accommodait donc de notre anthropocentrisme, voire même d'un centrisme racial et social d'une certaine anthropologie biologique de la fin du XIX<sup>e</sup> siècle. Scalisme, anthropocentrisme, maintien de la tradition et du progrès, les systématiciens défenseurs des grades participaient donc parfois inconsciemment d'une confusion en systématique entre partage des attributs, recherche d'un ordre naturel et discours de valeurs.

### Généalogie et phylogénie

En toute rigueur, aujourd'hui, une généalogie est une figure dont les liens joignent des individus concrets, c'est-à-dire individuellement identifiés, par des relations d'ancêtres à descendants. Ceci est particulièrement légitime au plan empirique lorsque l'on dispose de registres d'état-civil et que les individus sont, effectivement, individuellement identifiés. Le problème pour les naturalistes, c'est que les ancêtres sont à jamais inconnus : ils ont définitivement disparu et n'ont pas laissé d'état-civil. Une caractéristique de la systématique éclectique jusque dans les années 1970 est qu'elle confondait les relations de nature généalogique (qui descend de qui) avec les relations phylogénétiques (qui est plus apparenté à qui que d'un troisième) en inventant des relations d'ancêtres à descendants entre taxons de haut rang et en assignant à des fossiles le statut d'ancêtres incarnés. Même si, le plus souvent et chez les plus avisés des systématiciens, on savait quoi penser de ces métaphores généalogiques, elles devinrent tout de même un programme de recherches qui cherchera à se justifier y compris sous la plume d'un Ernst Mayr ou d'un Alfred Romer jusqu'au milieu des années 1970. Faisons le point aujourd'hui :

- Dans un arbre généalogique théorique (tel que celui de Darwin, 1859), on exprime des relations d'ancêtres à descendants entre individus abstraits. Cette généalogie relève de l'ontologie : elle parle de ce qui est, ou plutôt, de ce qui a dû être.
- Dans une généalogie concrète expérimentalement approchée, on exprime des relations d'ancêtre à descendants entre individus concrets (et là, on remarquera que, précisément, on n'a plus besoin des algorithmes de reconstruction phylogénétique lorsque l'on mène une enquête génétique de paternité, ou que l'on consulte les registres d'état-civil. Lorsque l'identité des ancêtres est à portée de main, on n'a plus besoin de phylogénie).
- Dans une phylogénie telle qu'on la reconstruit après Hennig, on exprime d'abord «qui partage quoi avec qui» entre individus concrets grâce à un graphe connexe. Grâce au cadre théorique, ces partages sont interprétés en termes de degrés relatifs d'apparement. Une telle phylogénie relève de l'épistémologie car le partage des attributs dont on a maximisé la cohérence est *un moyen* par lequel on va approcher *la connaissance* d'une partie de la grande généalogie, tout au moins son reflet. Les relations d'ancêtre à descendants sont alors conçus, dans une phylogénie, entre *ancêtres abstraits* partiellement reconstitués et *descendants concrets*, ou bien entre ancêtres abstraits et ancêtres abstraits lorsque l'on s'intéresse au lien entre deux branches internes.

L'arbre phylogénétique du vivant, résultat concret d'une enquête scientifique partielle menée sur un échantillon du vivant, est souvent confondu avec l'arbre généalogique du vivant qui est, lui, théorique. Il représente ce qui est, ou plutôt ce qui a été : une grande généalogie constituée de milliards d'ancêtres et de descendants reliés entre eux par des relations génétiques et à jamais inconnus individuellement. À travers eux, les processus de l'évolution se sont déroulés dans le sens du temps, du passé vers le présent. Cet arbre de la vie est une abstraction ontologique autant qu'une nécessité théorique qui explique le partage d'attributs communs par des espèces qui ne se croisent pas entre elles aujourd'hui.

Mais il faut insister sur l'importance de ne pas confondre généalogie (ce qui est) et phylogénie (les moyens d'accéder incomplètement à cette dernière : notamment nous n'avons pas les ancêtres concrets). Cette grande généalogie appartenant au passé restant théorique, nous ne pouvons qu'en inférer des parties par une démarche rétrospective : c'est à partir des états de caractères partagés par des organismes d'aujourd'hui que nous inférons ce qu'ont dû porter certains de leurs ancêtres communs, ancêtres qui resteront des portraits-robots ; nous construisons ainsi des arbres phylogénétiques. La reconstruction phylogénétique met en évidence des conditions du passé (états de caractères chez des ancêtres) qui ont dû donner naissance à l'état actuel du monde (états de ces caractères chez les descendants actuels). En principe, au plan pédagogique, si l'on veut respecter la signification scientifique de ces arbres, seul l'arbre généalogique théorique peut être parcouru du passé vers le présent (comme le faisait DARWIN 1859 dans son chapitre IV). L'arbre phylogénétique devrait être parcouru du présent vers le passé (comme le font DAWKINS, 2007 ou VAN WAEREBEKE *et al.*, 2008).

### Les ancêtres sont-ils connaissables ?

Il peut sembler paradoxal que, se réclamant de Darwin, la systématique phylogénétique moderne ait pu tant augmenter son pouvoir d'investigation en déclarant les ancêtres inconnaisables concrètement, tout en fondant son programme de classification sur la filiation.

La généalogie pure disconvient au projet taxonomique, comme le signale DUPUIS (1988) avec force :

*L'image de l'arbre généalogique est la seule que, malgré sa popularité, j'aurais aimé passer totalement sous silence. Abstraction faite de la fantaisie qui a longtemps régné en ce domaine, ma réticence a un motif très sérieux. Toutes les représentations d'une taxinomie sous forme d'arbres véritables, avec troncs, branches et feuilles, comme dans les planches de Haeckel (1866, t. II), sont trompeuses. Un phylum, qui naît probablement comme une infime petite pousse, n'est pas un tronc mais un ensemble de feuilles ! À cultiver à la lettre la métaphore de l'arbre généalogique, on*

*en vient à croire que le phylum a précédé les classes, qui auraient précédé les ordres, qui auraient précédé les espèces, et à croire aussi qu'un phylum vient d'un phylum, une classe d'une classe, etc. Je déconseille tout à fait l'image de l'arbre généalogique.*

DUPUIS expose ses réticences dans un contexte où il s'agit de cadrer le projet taxonomique. Pour le classificateur, la relation d'ancêtre à descendant n'est pas opérationnelle, elle est même encombrante tant elle nous pousserait à abandonner le nominalisme qui sous-tend tout le projet de la systématique en confondant nos concepts classificatoires (les taxons et leurs rangs, c'est nous qui les créons) avec la dynamique évolutive elle-même, c'est-à-dire en attribuant une existence évolutive réelle à nos concepts taxonomiques, en attribuant des relations génétiques d'ancêtres à descendants entre taxons de haut rang, ce qu'on fit sous le règne de la systématique éclectique avec le peu de rigueur que déplore DUPUIS : les mammifères «descendent» des reptiles, les «reptiles» descendent des amphibiens... Il ne s'agit donc que d'exclure les ancêtres *concrets* du champ méthodologique de la classification, tout en sachant que des ancêtres *abstraites* restent *requis sur le plan théorique*. Mais faut-il pour autant que les ancêtres *concrets* soient *exclus de la phylogénie* ? La réponse est oui, puisque la phylogénie moderne a justement été créée pour établir des relations d'affiliation entre individus *concrets* lorsque leurs ancêtres ont *définitivement disparu*. En revanche, DAYRAT (2005) répond que non, avec deux défauts : il ne distingue pas dans son article les ancêtres *abstraites* et les ancêtres *concrets*, préconisant des méthodes qui identifient des ancêtres *concrets*. Il ne distingue pas le niveau théorique du niveau empirique : il y a la généalogie théoriquement requise d'une part (dont parlait Darwin), et les possibilités empiriques de l'approcher d'autre part (ce dont parle Hennig).

La reconstruction phylogénétique moderne a fait un bond en avant avec HENNIG (1950, 1966) et ses successeurs en déclarant les ancêtres *inconnaisables*. DAYRAT (2005) identifie deux causes à l'exclusion des ancêtres (*concrets*) de la démarche phylogénétique. La première résiderait dans le fait que la relation ancêtre à descendants aurait été

déclarée comme non testable au sens poppérien du terme par les cladistes. Nous pouvons évacuer assez vite cette première cause. Certes, les cladistes de la fin des années 1970 ont eu des prétentions poppériennes assez naïves, mais une autre épistémologie peut être proposée pour la cladistique (DELEPORTE & LECOINTRE, 2005) et ce débat ne détermine ni n'épuise la question de l'accès aux ancêtres. La seconde raison résiderait dans le fait que Hennig était avant tout un taxonomiste et que son programme fut de faire de la phylogénie *pour* faire de la classification. Si toute l'information d'une série d'ensembles emboîtés (diagramme de Venn) est contenue dans un arbre phylogénétique, à l'inverse toute l'information contenue dans un arbre phylogénétique n'est pas dans le diagramme de Venn. Les relations d'ancêtres à descendants, précisément, ne figurent pas dans le diagramme de Venn, alors qu'elles sous-tendent théoriquement l'arbre phylogénétique (si l'on admet que dans un arbre phylogénétique il y a *au moins* des relations *indirectes* d'ancêtres à descendants entre des ancêtres *abstrait*s et des descendants). Hennig se serait comporté vis-à-vis de l'arbre phylogénétique comme il aurait convenu de se comporter à l'égard d'un graphe connexe non cyclique, c'est-à-dire un « arbre » dépourvu de sa justification théorique et de son sens biologique. Effectivement, si toute l'information d'une série d'ensembles emboîtés (diagramme de Venn) est contenue dans le graphe connexe non cyclique lui correspondant, à l'inverse toute l'information contenue dans ce graphe connexe non cyclique est effectivement contenue dans le diagramme de Venn. Cette seconde raison est recevable. Mais DAYRAT (2005) semble revenir à une confusion entre le niveau ontologique (celui de la généalogie théorique et ses relations d'ancêtres à descendants) et le niveau épistémologique (celui de la phylogénie où seules les relations de groupes-frères entre entités concrètes sont possibles, et où les ancêtres sont abstraits) qu'avait su éviter Hennig. Il ne va pas au fond du problème. Les méthodes qu'il préconise n'identifient pas, de toute façon, des relations directes d'ancêtres à descendants entre individus concrets, mais des conjectures de relations indirectes d'ancêtres à descendants entre concepts (un fossile

unique ou même un lot de fossiles identiques qui ont valeur d'espèce, par exemple). Ces concepts ne peuvent être que des concepts classificatoires. On reste dans la légitimité du classificateur qui était celle de Hennig.

DAYRAT (2005) justifie sa position en rappelant ce que préconisait DARWIN (1859) en matière de classification. Certes, DARWIN (1859) considérait certainement qu'un arbre généalogique était préférable à une classification pour représenter les relations entre des formes éteintes et actuelles, et il va même jusqu'à souhaiter que ces classifications deviennent elles-mêmes des arbres généalogiques, mais DUPUIS (1986) souligne que c'était là un idéal pour lui hors d'atteinte. On peut être surpris que DAYRAT n'ait pas remarqué qu'en matière de classification, Darwin se résout quand même à exprimer la recommandation de monophylie des groupes, même s'il n'utilise pas ce mot-là ni ne dit comment faire, justement parce que son discours reste au plan théorique et programmatique. DAYRAT omet de dire que chaque fois que Darwin parle de « généalogie », il parle d'une généalogie théorique et non d'une généalogie concrètement approchée, dont il dit par ailleurs que nous n'avons pas les registres d'état-civil. L'absence de moyens d'identifier, parmi les individus concrets, actuels ou fossiles, des ancêtres génétiques, nous conduit par la force des choses à ne concevoir ces ancêtres que comme des abstractions. Alors que DARWIN (1859) ne parlait que d'ancêtres abstraits, DAYRAT (2005) préconise des méthodes qui, grâce aux informations supplémentaires pourvues par la stratigraphie, assignent à des fossiles concrets le statut d'ancêtre possible. Les méthodes dont parle DAYRAT ne produisent que des *conjectures d'ancestralité*. La véritable relation génétique qui convient aux véritables généalogies reste *indémontrable* à ces temps de divergence. La relation d'ancêtre à descendant la plus pure, c'est-à-dire entre deux individus concrets (et non entre deux concepts classificatoires), reste inaccessible empiriquement, quoiqu'en fassent les méthodes montrées en exemple par DAYRAT (2005).

On ne peut raisonnablement que maintenir une opposition aux conclusions de DAYRAT (2005) :

«*Les relations d'ancêtres à descendants devraient être étudiées aussi souvent que possible parce qu'elles sont des représentations plus précises de l'histoire évolutive que ne le sont les relations de groupes-frères*». Si nous sommes à une échelle populationnelle intra-spécifique et à des temps de divergence très courts, peut-être... mais là le projet phylogénétique n'est plus convoqué ; d'autres outils sont disponibles. Si des relations d'ancêtres à descendants sont conçues entre autre chose que des individus concrets dont on peut faire preuve du lignage génétique, alors il ne peut s'agir que d'une relation génétique entre deux taxons dont celui qui donne naissance à l'autre est paraphylétique. DAYRAT (2005) insiste sur le fait que la recherche de relations d'ancêtres à descendants reste un programme de recherche darwinien. Certes, Darwin montre dans la seule figure de *L'Origine* une généalogie et en parle comme tel. Mais cela ne rend pas pour autant des ancêtres *concrets* accessibles, ni même la preuve génétique directe ou indirecte d'un lignage généalogique, même si l'on a pour programme de faire autre chose que de la classification. Les méthodes actuelles qui sont citées par DAYRAT (2005), en prétendant identifier des ancêtres concrets, sont sans doute des méthodes darwiniennes... mais elles ne sont pas des méthodes phylogénétiques.

Pour conclure, la filiation sous-tend toute la logique phylogénétique depuis HENNIG (1950) et même le programme classificatoire depuis DARWIN (1859), même s'il a fallu cent ans pour que le programme darwinien en cette matière puisse être réalisé. Darwin a exprimé dès la première édition de *L'Origine des espèces* la recommandation de monophylie des taxons sans utiliser le mot, et surtout sans dire comment trouver, concrètement, les groupes monophylétiques. La phylogénétique moderne se nourrit d'un paradoxe (TASSY 1994) qui n'est qu'apparent : tout le travail est sous-tendu par l'idée de filiation, mais les ancêtres sont déclarés inconnaisables. En réalité, il ne s'agit que d'une lucidité méthodologique : entre espèces, les individus *concrets* ancêtres au sens génétique du terme sont *inconnaisables empiriquement*, mais les ancêtres *abstraits* restent *indispensables théoriquement*.

## INTERROGER LA PHYLOGÉNIE

La définition de la phylogénie comporte trois difficultés. Premièrement, sa définition d'aujourd'hui ne ressemble pas à la définition qu'en ont donné deux grands noms, les premiers à l'avoir utilisée, à savoir HAECKEL (1866), qui inventa le mot, et DARWIN (1872) dans ses cinquième et sixième éditions de *L'Origine des espèces*. Deuxièmement, il est difficile de trouver une définition moderne du mot «phylogénie» dans les sources francophones. Par exemple, dans le très consulté *Reconstruction phylogénétique* de DARLU & TASSY (1993), passées les premières pages où l'on aborde les définitions historiques du terme, il faut attendre la page 36 pour apprendre que la phylogénie, c'est le cladogramme auquel on a adjoint l'échelle des temps ou des longueurs inégales de branches en proportion du nombre de synapomorphies ou d'autapomorphies qu'elles portent. Troisièmement, d'un auteur à l'autre on trouve pour la phylogénie soit des définitions qui relèvent de l'ontologie et qui tendent tacitement à confondre les relations d'ancêtres à descendants et les degrés relatifs d'apparementement (chez HAECKEL, 1866 ; chez EDWARDS, 2009), soit des définitions qui font relever la phylogénie de la seule épistémologie (chez TASSY & BARRIEL, 1995 ; ou LECOINTRE, 2008) et qui distinguent clairement les deux. Dans ce dernier cas la phylogénie n'est que le moyen indirect d'approcher la GT : à défaut d'ancêtres concrets on se contentera de la mise en évidence de degrés relatifs d'apparementement où les ancêtres sont des reconstructions abstraites obtenues par la mise en évidence des homologies secondaires (de PINNA, 1991).

Depuis HAECKEL (1866), nous retenons usuellement la définition floue de phylogénie comme «le cours historique de la descendance des êtres organisés» (DARLU & TASSY, 1993), mélange d'une définition haeckelienne qui voulait parler de l'enchaînement des formes animales et végétales au cours du temps, et d'une définition darwinienne parlant des «lignes généalogiques de tous les êtres organisés». Cependant, le jeu de miroir entre Darwin et Haeckel repose sur un malentendu (DAYRAT, 2003), nous l'avons vu.

Avec HENNIG (1966), ce qui est théoriquement requis (la généalogie passée) et ce qui est pratiquement possible (quel type de relations sommes nous vraiment capables de montrer ?) sont deux choses clairement distinguées. Se superpose à cela une qualité complémentaire à celle de Darwin. Si Darwin travailla plus à éclairer les processus par lesquels les espèces se transforment qu'à les classer (bien qu'il ait mené une réflexion sur la classification dans son chapitre XIII, cf. DUPUIS, 1986), Hennig va s'attacher davantage à fonder les bases d'une bonne classification selon le programme darwinien (DAYRAT, 2005). Tout se passe comme si Hennig était le classificateur qui manqua à Darwin (cf. DUPUIS, 1978, 1986, 1992).

Examinons, dans la continuité de Hennig, une définition épistémologique de la phylogénie. La phylogénie n'a pas besoin d'une définition ontologique. En effet, l'arbre généalogique des générations successives d'êtres vivants suffit comme base ontologie au travail de l'évolutionniste et du systématicien. Parce que l'histoire des gènes ne suit pas nécessairement celle des espèces (DOYLE, 1997; MADDISON, 1997), la définition d'EDWARDS (2009) propose de réserver le terme de « phylogénie » au seul arbre des espèces et non aux arbres des gènes (qui sont, la plupart du temps, supposés être témoins de l'histoire des espèces). Cette proposition est insuffisante. Si par « arbre » on entend le GCNC, on ne construit jamais qu'un arbre sur la base de caractères choisis. Une synthèse de multiples arbres de gènes dessinée afin de représenter un arbre d'espèces restera sur le plan technique l'arbre fourni par *ces gènes là*, fussent-ils plusieurs : il s'agira toujours seulement d'une phylogénie tirée de parties d'organismes. Dès lors, réserver le terme de phylogénie au seul arbre des espèces revient à transférer le terme dans le registre de l'ontologie et confondre celle-ci avec la généalogie théorique du vivant. En d'autres termes, sur le plan épistémologique, c'est-à-dire sur le plan des moyens mis en œuvre pour reconstituer partiellement des degrés relatifs d'apparement entre entités, il n'y aura jamais de phylogénies d'espèces, mais seulement des phylogénies individuelles de gènes (même si celles-ci coalescent et racontent la même

histoire). Non, la phylogénie ne doit pas être définie en rapport à la nature des entités dont on cherche les relations (d'espèces, de gènes, de langues, de voies métaboliques, etc.), ce que cherche à faire EDWARDS (2009). Elle a, au contraire, tout intérêt à bénéficier d'une définition qui spécifie sa démarche propre et qui garantisse sa lucidité méthodologique, en somme une définition épistémologique où l'on ne confond pas les relations d'ancêtres à descendants (empiriquement inaccessibles mais requis sur les plans ontologique et théorique) et les degrés relatifs d'apparement, seule relation que nous puissions mettre en évidence. Selon cette définition, un arbre peut être qualifié de phylogénétique s'il suit les deux règles suivantes :

1. Le concept de descendance avec modification a été formulé dans la méthode qui a présidé à sa construction. Ce n'est pas seulement dans la polarisation des caractères que tient le concept de descendance avec modification, mais aussi et surtout dans le projet hiérarchique. La raison pour laquelle une classification biologique est une hiérarchie présentée sous forme arborescente est la descendance avec modification (quelle que soit la nature des entités qui y souscrivent : langues, voies métaboliques, espèces, gènes, etc.). Cependant, ce point a été âprement discuté et les interprétations divergent (cf. BROWER, 2000 ; RIEPPEL, 2005).
2. Selon TASSY & BARRIEL (1995), l'arbre est phylogénétique s'il permet une découverte *a posteriori* des homologies et des homoplasies. En particulier, la méthode de parcimonie nous permet une découverte de nos succès (homologies secondaires de PINNA 1991) et de nos erreurs (homoplasies) dans ce que nous avons supposé initialement comme caractères homologues (homologies primaires). Cette découverte permet de déduire l'état de chaque caractère chez chaque ancêtre hypothétique, aux nœuds de l'arbre (ce que permettent également, en principe, les méthodes de vraisemblance). Plus précisément, ce sont ces réponses sur les homologies qui permettent d'inférer des parties d'ancêtres, et ceci est une excellente raison de souscrire à la définition de

TASSY et BARRIEL (1995). De plus, l'algorithme de WAGNER (1961) utilisé dans les méthodes de parcimonie contemporaines (DARLU & TASSY, 1993 : 79) maximisent ce que FARRIS (1979, 1983) appelle l'«*explanatory power*», c'est-à-dire maximisent la contiguïté des états de caractères identiques, et donc l'*explication* de ces états par une ascendance commune, donc maximisent l'information phylogénétique sur les caractères (TASSY, 1994 : 107-108), contenu informatif que FARRIS démontre supérieur dans un cladogramme que dans un phénogramme. En d'autres termes, les arbres qui ne supposent pas un maximum d'états de caractères expliqués par l'ascendance commune alors qu'ils pourraient le faire sont de mauvais arbres, ou dit autrement, des arbres non optimaux.

Selon cette définition, les méthodes de distances sont incomplètes et sont qualifiées de «*pseudo-phylogénies*» par TASSY & BARRIEL (1995). Les principaux arguments que l'on peut généralement dégager sont les suivants.

1. Les arbres de distances ne permettent pas d'inférence *in fine* sur l'homologie des caractères, puisque l'on ne travaille pas en prise directe avec leurs états. En fait, comme l'écrivent DARLU & TASSY (1993 : 81), c'est cette étape d'estimation des états ancestraux qui constitue toute la différence entre procédures de parcimonie et analyses de distances. Chez celles-ci, une fois l'alignement fait, l'homologie primaire des états est dès le départ réduite sous forme de distances et les paris qu'elle contient resteront sans résolution.

2. Le plaquage sur l'arbre de distances des états de caractères ne peut être admis comme exercice d'inférence de leur homologie secondaire. En effet, les méthodes de distances peuvent produire des regroupements sur la base de simplésiomorphies, comme montré dans LECLERC *et al.* (1998). Quel sens aurait la découverte des homologies/homoplasies sur une topologie dont les artefacts sous-jacents sont incompatibles avec la notion de synapomorphie? Cela reviendrait à nier l'intérêt de l'apport de Hennig et à mélanger des méthodes aux propriétés

de restitution des données dans l'arbre inégales (FARRIS, 1979). Il est bien plus cohérent de plaquer directement ces caractères sur un arbre issu d'une méthode dont c'est la vocation même, c'est-à-dire qui regroupe sur la base de synapomorphies.

L'arbre phylogénétique est donc, pour TASSY & BARRIEL (1995), celui qui autorise l'identification du couple homologie/homoplasie. La nécessité de cette définition épistémologique, qui vient en partie du fait que les taxons sont définis par des caractères homologues (PATTERSON, 1982, 1988 ; NELSON, 1994), est cruciale pour le projet du systématique (le plus souvent morphologiste) dont le travail est de créer des taxons monophylétiques, donc d'identifier des homologies secondaires. Elle semble cependant parfois superflue pour le généticien (EDWARDS [2009] ne l'envisage même pas !), qui travaille sur des caractères bien moins complexes, et donc individuellement moins dignes d'intérêt, et qui aurait tendance à prendre n'importe quelle arborescence pour phylogénétique – d'où la grande diversité des méthodes de construction d'arbres utilisés en phylogénie moléculaire. À ce sujet, TASSY & BARRIEL (1995) se montrent intransigeants, en ne réservant l'appellation de «*phylogénétique*» qu'aux strictes méthodes de parcimonie : «*L'analyse cladistique et a fortiori l'analyse de parcimonie sont souvent considérés comme des méthodes parmi d'autres. Nous prétendons qu'elles sont une seule et même méthode qui est la méthode phylogénétique*» (l'insistance est des auteurs).

Cependant, les méthodes probabilistes donnent cette possibilité de placer les homologies aux nœuds. Les méthodes probabilistes sont donc phylogénétiques :

*Toutes les méthodes qui tendent à esquiver l'homologie – parce que c'est un concept, opératoire mais faillible, parce que c'est un problème biologique, parce qu'elle reste du domaine de l'hypothèse – ne peuvent prétendre être du domaine de la phylogénétique au sens de KIRIAKOFF (1963) : «*la science des constructions phylogénétiques*». Notre conclusion sera qu'il y a à la fois filiation et transformation depuis l'arbre phylogénétique jusqu'au cladogramme et que seulement deux des méthodes informatiques en cours actuellement*



*répondent aux critères de la phylogénétique : la méthode dite de parcimonie et, dans une certaine mesure, celle dite de maximum de vraisemblance.*

TASSY & BARRIEL (1995) en conviennent, mais du bout des lèvres. On reproche couramment aux méthodes probabilistes de ne pas replacer les homologues aux nœuds alors qu'elles pourraient le faire. Ce reproche ne doit pas être fait aux méthodes probabilistes elles-mêmes, mais au type de caractères auxquels on les applique aujourd'hui, et à l'intérêt qu'on leur porte. En effet, les méthodes probabilistes sont majoritairement appliquées à la comparaison de séquences génétiques pour produire des « phylogénies moléculaires » aux caractères desquelles le systématicien s'intéresse peu le plus souvent, il faut bien le dire. Que le nucléotide n° 727 de tel gène soit une adénine homologue chez tous les Proboscidiens n'a pas la même portée, le même piment scientifique, la même complexité que s'il s'agit d'une homologie trouvée dans l'ouverture de l'orbite dans le maxillaire ou dans les os qui entourent le trou auditif externe. Ceux des systématiciens qui sont un peu biochimistes, qui, connaissant quelque peu les enjeux fonctionnels d'un changement de nucléotide à une position donnée d'un gène donné, tendent à raisonner autrement et à s'intéresser aux conséquences évolutives d'un changement moléculaire. Enfin, les méthodes probabilistes sont récusées par TASSY & BARRIEL (1995) parce qu'elles nécessitent un modèle en amont :

*« Quant au choix des modèles, c'est là un débat sans fin. Faut-il connaître le processus évolutif pour construire la phylogénie ou a-t-on besoin d'une phylogénie pour découvrir le processus évolutif responsable de cette phylogénie ? Les cladistes ont choisi depuis longtemps de répondre positivement à la deuxième question (ELDREDGE & CRACRAFT, 1980 ; NELSON & PLATNICK, 1981) ».*

Mais cette alternative est surfaite. Elle assigne aux modèles des défauts qui ne sont pas les leurs. Elle accuse les modèles tenant compte de « processus » évolutifs du défaut de précédence, parce que ces mêmes précédences de processus prétendument connus *a priori* s'accompagnaient, dans les pratiques passées de la systématique éclectique, d'un manque

criant de formalisation. Cependant, les modèles sont précisément un effort de clarification et de formalisation de ce que l'on prétend connaître. D'autre part, cette précédence est vécue comme circulaire en raison d'une vision réfutationniste poppérienne et hypothético-déductive de la systématique. Mais depuis, d'autres épistémologies ont été proposées pour la systématique (RIEPPPEL, 2003), certainement plus abductives qu'hypothético-déductives. Le « cohérentisme » des cladistes peut alors se muer en « fondhérentisme » (le terme est de HAACK, 2000 ; voir RIEPPPEL, 2005) où les connaissances de fond sont prises en compte pourvu qu'elles soient explicites (LECOINTRE & DELEPORTE, 2005 ; DELEPORTE & LECOINTRE, 2005). L'interaction des connaissances sur les processus avec l'arbre n'est pas circulaire : connaissances de fond (relevant des structures comme des processus) et nouvelles données progressent ensemble, par une nouvelle étape de mise en cohérence, vers de nouvelles connaissances. Pour reprendre la métaphore fondhérentiste (RIEPPPEL, 2005) :

*« [...] le schéma est celui d'une grille de mots croisés dans laquelle on introduit de nouveaux mots, mais d'une manière qui doit être cohérente avec tous les autres mots qui sont déjà en place, et où les vieux mots déjà en place sont susceptibles d'être révisés à la lumière des nouveaux mots qui sont introduits. Le fondhérentisme cherche à combiner l'intégration explicative avec l'ancrage expérientiel. »*

Ensuite, cette alternative résulte plus de l'élaboration d'une règle idéale et normative du raisonnement que d'une prise en compte des contraintes réelles inhérentes aux pratiques de la systématique (c'était d'ailleurs aussi le problème de Karl Popper à l'égard de la démarche scientifique). Prise à la lettre, elle conduit à l'impossibilité même de construire une matrice de caractères anatomiques. En effet, elle conduit logiquement à reprocher à un anatomiste d'omettre un caractère de sa matrice parce que, ayant observé ce caractère au long de séries de développement embryonnaire, le processus de développement montre qu'il résulte de processus ontogénétiques différents entre les taxons, et qu'il n'est donc pas homologue. Ces cas de figure ne sont pas rares, par exemple chez les vertébrés depuis que les techniques

d'éclaircissement des tissus et de coloration d'os et de cartilages permettent d'observer la mise en place des structures squelettiques le long d'une série d'embryons à des stades rapprochés. Faudrait-il ignorer cette connaissance et inclure quand même le caractère dans la matrice au motif que le développement embryonnaire est qualifié de processus (mise en évidence de relations de cause à effet)? Mais il y a encore plus général : toute appréhension d'un caractère par un investigateur est forcément chargée de connaissances, et investie théoriquement? (RIEPPPEL, 2005). Parmi ces connaissances sont incluses celles sur les processus biologiques. Si le systématique était conscient de *tous* ses postulats et qu'il devait en exclure tout ceux qui dépendent pour tout ou partie de connaissances sur les processus, il ne pourrait tout simplement pas énoncer la moindre hypothèse d'homologie primaire. Lucidité pour lucidité, une épistémologie fondhérentiste consciente est préférable à une pureté cladiste (version 1980) épistémologiquement illusoire et empiriquement impossible à tenir.

Enfin, l'alternative est étonnante lorsque des systématiques parlent de « découvrir des processus ». Les systématiques n'ont pas pour vocation de découvrir des processus. Ils ont pour vocation professionnelle de gérer de manière cohérente les rapports entre les concepts classificatoires, les mots et les choses. Ici un autre débat s'ouvre, celui des horizons d'investigation qui sont ceux de chaque spécialité en biologie ; autrement dit, la nature des questions auxquelles nous répondons dans nos articles. Une phylogénie peut conduire à faire des classifications. Une phylogénie peut peut-être conduire à découvrir des processus (et même cela il faudrait le discuter : une phylogénie peut-elle vraiment prouver une relation de cause à effet?). Mais ce ne sont généralement pas les mêmes communautés professionnelles qui atteignent ces objectifs. Et si ces deux approches sont menées dans une seule et même tête, ce n'est généralement pas au même moment, au moins parce que nous avons besoin qu'une phylogénie soit fiable avant de commencer à élaborer d'autres types de connaissances à partie d'elle... Et elle est rarement fiable dans son entièreté à partir d'un seul jeu de données (tout

systématique sait qu'un arbre a toujours, à côté de ses parties solides, des parties fragiles). Pour ne prendre qu'un exemple, les relations d'apparement à grande échelle entre les téléostéens sont si difficiles à élucider, et les classifications classiques si confuses qu'on peut y épuiser plusieurs carrières sans avoir besoin de poser une seule question en termes de processus. Certes, les processus de l'évolution peuvent être pris en compte comme donnée pertinente utile à l'élaboration de réponses à un questionnement purement systématique, mais ce n'est pas pour autant que ces questions ultimes deviennent des questions relevant des processus de l'évolution. Dans une phylogénie générale de téléostéens, aucune relation de cause à effet n'est élucidée. Pourtant, il a fallu beaucoup d'énergie pour que cette phylogénie soit fiable. Il convient donc de souligner la différence qu'il peut y avoir dans les rapports entre patrons et processus tout au long de la démarche d'investigation d'un biologiste. Contrairement à l'alternative qui est proposée ci-dessus, tous les biologistes, quels qu'ils soient, ont besoin comme données sources et comme connaissances de fond à la fois des données et connaissances relevant des patrons, et des données et connaissances relevant des processus. En revanche, c'est le type de question qu'ils posent – le point ultime de leur investigation – qui détermine s'ils sont dans une science de patron ou une science de processus, parce que la question va déterminer le type de preuve qui est à l'œuvre. L'alternative proposée ci-dessus est artificielle et présente le défaut principal de ne pas distinguer ce qui relève des données et connaissances sources de ce qui relève de la nature de la question posée. Des auteurs qui écrivent sur les méthodes de la systématique prescrivent une abstinence en matière de processus en amont de la construction de l'arbre au motif que cet arbre pourrait amener des connaissances sur les processus qu'eux-mêmes ne questionnent même pas... s'ils sont de vrais systématiques ! À moins que par « processus » soit entendue la simple mise en série des états d'un même caractère comme possiblement dérivés les uns des autres, que permet effectivement l'arbre et lui seul. Ici le terme de « processus » serait alors mal choisi, car il n'y a pas démonstration expérimentale

d'une relation de cause à effet. Il n'y a que mise en cohérence maximale d'attributs, mise en preuve qui ne relève que des sciences des patrons.

Pour finir sur cette alternative présentée par TASSY & BARRIEL (1995), et si on élargit le champ, ce qui est souhaitable, les auteurs qui ont écrit ces lignes entendent sans doute leur donner une portée plus générale. Ils prescrivent une abstinence en matière de processus en amont de la construction de l'arbre, au motif que cet arbre pourrait amener des connaissances sur les processus qui pourraient être élaborées par d'autres métiers que celui de la systématique. Si tel est le sens donné à cette prescription, elle se justifie alors plus par son idéalisme que par son réalisme au regard des conditions de la recherche. D'autres chercheurs, avec d'autres types de questions en tête que celles des systématiciens, auront besoin d'une phylogénie fiable. Cette dernière sera rarement celle d'un seul auteur, parce que les phylogénies publiées par les laboratoires sont rarement fiables dans leur entièreté. C'est la confrontation de diverses phylogénies obtenues par différentes équipes et différentes sources de données sur les mêmes taxons (ou presque) qui établit la fiabilité d'une hypothèse phylogénétique. Dans ce processus de validation des connaissances objectives, on constatera que certaines équipes auront suivi la prescription pour construire leur arbre, d'autres non ; et la corroboration des mêmes clades par des sources différentes aura un poids bien plus fort pour remporter l'assentiment que le respect d'une abstinence initiale au regard des « processus ». Dotée d'une portée aussi générale, la prescription des cladistes de cette époque du début des années 1980 confond la phase de démonstration avec la phase de validation des connaissances produites.

Terminons-en avec la question de savoir quels arbres sont phylogénétiques et quels arbres ne le sont pas. Un graphe connexe non cyclique (ce qu'on appelle communément aujourd'hui un arbre), sur le plus pur plan algorithmique, ne montre que les partages d'attributs : « qui partage quoi avec qui ». Il ne deviendra phylogénétique que nourri de la théorie de l'évolution sous son bagage minimal du « *descent*

*with modification* ». De par sa définition purement épistémologique, une phylogénie ne donne en première instance que des relations de groupes-frères entre entités concrètes, pas des relations d'ancêtres à descendants entre entités concrètes. La méthode phylogénétique n'est *pas faite pour cela*, tout simplement. Nous ne sommes pas capables de prouver empiriquement une véritable relation génétique d'ancêtres à descendants entre deux individus concrets d'espèces différentes, tout simplement (TASSY, 1994). L'outil phylogénétique hennigien est précisément fait pour les cas de figure où les ancêtres de ces individus concrets sont inconnaissables concrètement. Lorsque nous sommes susceptibles de prouver génétiquement une relation d'ancêtre à descendant, au sein d'une même population concrète d'une même espèce concrète, et dans des temps très courts, de l'ordre de quelques générations, alors ce ne sont plus les outils de la phylogénie qui sont utilisés, mais ceux de la génétique des populations. Cependant, il subsiste, dans la phylogénie nourrie de son fond théorique, des relations d'ancêtres à descendants. Mais celles-ci se conçoivent entre un ancêtre *abstrait* (partiellement reconstitué) et un descendant concret s'il s'agit d'une branche terminale, ou bien entre un ancêtre abstrait et un descendant abstrait s'il s'agit d'une branche interne. Cette relation d'ancêtre à descendant n'est pas directe comme elle peut l'être dans une généalogie concrète, dans une phylogénie, elle est *indirecte*. Cela signifie qu'entre un ancêtre abstrait et son descendant, beaucoup d'individus hypothétiques et non reconstitués, généalogiquement reliés, sont théoriquement supposés.

En somme, est phylogénétique l'arbre qui permet de restituer la réponse aux paris qui avaient été faits sur les homologues. En cela, les méthodes de distances ne sont pas des phylogénies. Elles ne sont, éventuellement, que des reflets de phylogénie, et encore, avec des postulats surnuméraires : comme l'arbre de distances ne montre que des relations de proximité en similitude globale (et non en apparemment), le « qui ressemble à qui » ne peut refléter le « qui est plus apparenté à qui » uniquement lorsque l'accumulation de la similitude globale est proportionnelle au degré d'apparemment. C'est au

prix de ce pari qu'un phénogramme ressemblera à une phylogénie. Bien des articles de phylogénie moléculaire font tacitement ce pari, aujourd'hui. Il est même courant que ce pari ne soit même plus conscient, et qu'on qualifie à tout va n'importe quel arbre de «phylogénie» sans plus s'intéresser à la méthode qui l'a construit.

### **CONCLUSION : INTERROGER LA CONCORDANCE ENTRE LES GRAPHES ET L'ARBRE GÉNÉALOGIQUE THÉORIQUE**

En distinguant la métaphore de l'arbre (niveau métaphorique souvent utilisé lors de la synthèse des connaissances), la généalogie théorique (niveau théorique : GTD ou GTR), le graphe connexe (niveau technique : GCNC ou GCC) et la phylogénie (niveau épistémologique), nous pouvons envisager sereinement les relations entre les trois dernières entités en cas de transferts horizontaux dominants (fig. 1). Au niveau théorique (fig. 1), tant que les transferts sont rares (GTD), l'utilité du GCNC pour construire des phylogénies en cas de GTD postulée est indiscutable. Lorsque des transferts horizontaux ou des hybridations sont à supposer au niveau théorique (GTC), l'analyse phylogénétique séparée de gènes ou de blocs de gènes via des GCNC multiples reste possible. La phylogénie se complexifie, elle doit alors, à l'étape de la synthèse et pour résumer les informations surnuméraires, incorporer des flèches d'une branche à l'autre et dans tous les cas indiquer de quoi il est question (quels gènes, quels traits... font l'objet d'un transfert). Mais dans ce cas on peut aussi, au niveau technique, passer directement au GCC comme outil d'investigation des relations si les transferts sont importants. Cela remet-il en cause la phylogénie? Aucunement. Parce que la phylogénie est définie au mieux sur le plan épistémologique, et non sur le plan ontologique : *la phylogénie est un graphe connexe qui permet d'exprimer des degrés relatifs d'apparentement et qui permet de trouver les solutions aux hypothèses d'homologies primaires qui avaient été initialement formulées dans la matrice.* Ces deux volets de sa définition impliquent les points

suiuants. La première partie de la définition implique que le graphe est théoriquement chargé du «*descent with modification*» (le «qui partage quoi avec qui» nous renseigne sur «qui est plus proche de qui», que le «descent» soit *vertical ou horizontal*) et que la phylogénie est précisément faite lorsque les ancêtres ne peuvent être concrètement identifiés. C'est pourquoi dans une phylogénie, il n'y a pas de relations d'ancêtres à descendants entre entités concrètes : les ancêtres sont théoriquement requis, et ils sont abstraits (si nous étions capables d'identifier des relations d'ancêtres à descendants entre entités concrètes, ce sont d'autres outils techniques qui seraient mobilisés). La seconde partie de la définition implique que les méthodes de distances, en exprimant par un GCNC des degrés de similitude globale, renoncent à l'homologie secondaire et ne peuvent être qualifiées de phylogénétiques. Enfin, cette définition de la phylogénie ne dépend pas de l'adéquation des GCNC ou des GCC soit à la GTD, soit à la GTR. L'adéquation des GCNC à des situations de GTD postulées n'est pas à discuter. En situation de GTR postulée, de multiples GCNC construits sur des jeux de caractères ayant subi le même «processus de discordance» (DOYLE, 1997; MADDISON, 1997) sont également envisageables. Enfin, les GCC peuvent être utilisés en cas de GTR postulée. Dans tout les cas, nous avons à faire à de la phylogénie car les trois cas de figure souscrivent aux éléments de définition fournis précédemment : nous avons des relations d'apparentement et les homologies secondaires sont finalement obtenues (tant qu'on dispose d'une racine). Il suffira de préciser phylogénie de «quoi» (des gènes, de portions de génomes, de populations, d'espèces...). On comprendra qu'une telle définition récuse la proposition d'EDWARDS (2009) de réserver le terme de «phylogénie» à une catégorie particulière d'entités biologiques (les espèces et les populations). En fait, il peut et il doit y avoir phylogénie de toute entité qui puisse souscrire au concept de «*descent with modification*». Ainsi, si le statut de la phylogénie n'est pas changé par un changement de forme de la GT, reste à savoir si la métaphore de l'arbre n'a pas été, et ne resterait pas aujourd'hui plus encombrante qu'utile, en raison de son imprécision.

## BIBLIOGRAPHIE

- AUGIER A., 1801. *Essai d'une nouvelle classification des végétaux conforme à l'ordre que la nature paraît avoir suivi dans le règne végétal*. Bruyset aîné et comp., Lyon.
- BARBANÇOIS-VILLEGONGIS C.H. DE, 1816. Observations sur la filiation des animaux, depuis le polype jusqu'au singe. *Journal de physique, de chimie, d'histoire naturelle et des arts*, 82 : 444-448.
- BEN HAMED M., 2009. La linguistique historique, nouveau terrain d'expérimentation de la phylogénie. In : T. Heams, P. Huneman, G. Lecointre & M. Silberstein (dir.), *Les mondes darwiniens – L'évolution de l'évolution*, p. 975-996, Syllepse, Paris.
- BROWER A., 2000. Evolution is not a necessary assumption of cladistics. *Cladistics*, 16 : 143-154.
- BUFFON G., 1755. *Histoire naturelle, générale et particulière avec la description du cabinet du roi*, tome cinquième. Imprimerie Royale, Paris. [Entrée : Table des chiens et de leurs variétés]
- BUFFON G., 1766. *Histoire naturelle, générale et particulière avec la description du cabinet du roi*, tome quatorzième. Imprimerie Royale, Paris. [Entrée : De la dégénération]
- CUNCHILLOS C. & LECOINTRE G., 2007. Ordering events of Biochemical evolution. *Biochimie* 89 : 555-573.
- DARLU P. & TASSY P., 1993. *Reconstruction phylogénétique. Concepts et méthodes*. [Collection Biologie théorique n° 7]. Masson, Paris.
- DARWIN C., 1859. *The Origin of Species*, First Edition. John Murray, Londres. Réédition, 1985 [Penguin Classics]. Penguin Books, Londres.
- DARWIN C., 1860. *The Origin of Species*, Second Edition. John Murray, Londres. Réédition, 1998 [Oxford World's Classics]. Oxford University Press, Oxford.
- DARWIN C., 1872. *The Origin of Species*, Sixth Edition. John Murray, Londres.
- DARWIN C., 1880. *L'origine des espèces au moyen de la sélection naturelle ou la lutte pour l'existence dans la nature traduit sur l'édition anglaise définitive* par Ed. Barbier. C. Reinwald, Paris (traduit par E. Barbier sur l'édition définitive (sixième édition) de « The Origin of Species », John Murray, Londres, 1872).
- DARWIN C., 1999. *La Filiation de l'Homme et la sélection liée au sexe précédé de L'anthropologie inattendue de Charles Darwin*. Traduction coordonnée par M. Prum. Syllepse, Paris. (Traduction française de « The Descent of Man » par M. Prum sous la direction de P. Tort)
- DARWIN E. 1794. *Zoonomia or the laws of organic life, second edition corrected (1796)*. J. Johnson in St Paul's churchyard, Londres.
- DAWKINS R., 2007. *Il était une fois nos ancêtres. Une histoire de l'évolution*. Robert Laffont, Paris.
- DAYRAT B., 2003. The roots of phylogeny : How did Haeckel build his trees? *Syst. Biol.*, 52 : 515-527.
- DAYRAT B., 2005. Ancestor-descendant relationships and the reconstruction of the tree of life. *Paleobiology*, 31 : 347-353.
- DENAMUR E., LECOINTRE G., DARLU P., TENAILLON O., ACQUAVIVA C., SADAYA C., SUNJEVARIC I., ROTHSTEIN R., ELION J., TADDEI F., RADMAN M. & MATIC I., 2000. Evolutionary implications of the frequent horizontal transfer of mismatch repair genes. *Cell*, 103 : 711-721.
- DELEPORTE P. & LECOINTRE G. (coords.), 2005. *Philosophie de la systématique* [Biosystema 24]. Société française de systématique. Paris.
- DOOLITTLE W.F. & BAPTESTE E., 2007. Pattern pluralism and the tree of life hypothesis. *Proc. Nat. Acad. Sci.* 104 : 2043-2049.
- DOYLE J.J., 1997. Trees within trees : genes and species, molecules and morphology. *Syst. Bot.*, 46 : 537-553.
- DUCHESNE A.N., 1866. *Histoire naturelle des fraisiers*. C.J. Panckoucke, Paris.
- DUPUIS C., 1978. Permanence et actualité de la systématique : la « systématique phylogénétique » de W. Hennig (Historique, discussion, choix de références). *Cahier des naturalistes*, 34 : 1-69.

- DUPUIS C., 1986. Darwin et les taxonomies d'aujourd'hui. In : P. Tassy (coord.), *L'ordre et la diversité du vivant*, p. 215-240, Fayard, Fondation Diderot, Paris.
- DUPUIS C., 1988. Le taxinomiste face aux catégories. *Cahier des naturalistes*, 44 : 49-109.
- DUPUIS C., 1992. Regards épistémologiques sur la taxinomie cladiste. Adresse à la XI<sup>e</sup> session de la Willi Hennig Society (Paris, 1992). *Cahier des naturalistes*, 48 : 29-56.
- EDWARDS S.V., 2009. Is a new and general theory of molecular systematics emerging? *Evolution*, 63 : 1-19.
- ELDREDGE N. & CRACRAFT J., 1980. *Phylogenetic patterns and the evolutionary process. Method and theory in comparative Biology*. Columbia University Press, New York.
- ESCOBAR-PARAMO P., SABBAGH A., DARLU P., PRADILON O., VAURY C., DENAMUR E. & LECOINTRE G., 2004. Decreasing the effects of horizontal gene transfer on bacterial phylogeny : the *Escherichia coli* case study. *Mol. Phyl. Evol.*, 30 : 243-250.
- FARRIS J.S., 1979. The informative content of the phylogenetic system. *Syst. Zool.*, 28 : 483-519.
- FARRIS J.S., 1983. The logical basis of phylogenetic analysis. In : N. Platnick & V.A. Funk (eds.), *Advances in cladistics*, Vol. 2, p. 1-36, Columbia University Press, New York.
- GAYON J., 1992. *Darwin et l'après Darwin*. Kimé, Paris.
- GAYON J., 2009. Mort ou persistance du darwinisme? Regard d'un épistémologue. *C.R. Palevol.*, 8 : 321-340.
- HAACK S., 2000. A foundherentist theory of empirical justification? In : E. Sosa & E. Kim, (eds.), *Epistemology, an anthology*, p. 226-236. Blackwell, Malden.
- HAECKEL E., 1866. *Generelle Morphologie der Organismen*. Reimer, Berlin.
- HAECKEL E., 1874. *Anthropogenie*. Engelmann, Leipzig.
- HENNIG W., 1950. *Grundzüge einer Theorie der phylogenetischen Systematik*. Deutscher Zentralverlag, Berlin.
- HENNIG W., 1966. *Phylogenetic Systematics*. Univ. of Illinois Press, Urbana.
- HUSON D.H. & BRYANT D., 2006. Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.*, 23 : 254-267.
- KIRIAKOFF S., 1963. Les fondements philosophiques de la systématique biologique. In : Centre national belge de recherche logique, Société belge de logique et de philosophie des sciences (eds.), *La classification dans les sciences*, p. 61-88, Duculot, Gembloux.
- LAMARCK J.B., 1809. *Philosophie zoologique*, 2 tomes. Dentu, Paris. Réédition en 1 volume, 1994. Flammarion, Paris.
- LAURIN M., 2008. *Systématique, paléontologie et biologie évolutive moderne : l'exemple de la sortie des eaux chez les vertébrés*. Ellipses, Paris.
- LECLERC M.C., BARRIEL V., LECOINTRE G., DE REVIERS B., 1998. Low divergence in rDNA ITS sequences among five species of *Fucus* (Phaeophyceae) suggests a very recent radiation. *J. Mol. Evol.*, 46 : 115-120.
- LECOINTRE G., 2007. Comprendre le matérialisme par son histoire. In : P. Charbonnat, *Histoire des philosophies matérialistes*, p. 13-33, Syllepse, Paris.
- LECOINTRE G. (dir.), 2008. *Comprendre et enseigner la classification du vivant*. Seconde édition. Belin, Paris.
- LECOINTRE G. (dir.), 2009. *Guide critique de l'évolution*. Belin, Paris.
- LECOINTRE G. & DELEPORTE P., 2005. Total evidence requires exclusion of phylogenetically misleading data. *Zoologica Scripta*, 34 : 101-117.
- LOPEZ P. & BAPTISTE E., 2009. Molecular phylogeny : reconstructing the forest. *C. R. Biologies*, 332 : 171-182.
- MADDISON W.P., 1997. Gene trees in species trees. *Syst. Biol.*, 46 : 523-536.
- MAILLET B. DE, 1755. *Telliamed*. Réédition, 1984. Fayard, Paris.

- MAUPERTUIS P.L. DE, 1751. *Essai de cosmologie*. Imp. Luzac, Leiden.
- MAYR E., 2004. *Après Darwin*. Quai des Sciences, Dunod, Paris.
- NELSON G.J., 1972. Comments on Hennig's «phylogenetic systematics» and its influence on ichthyology. *Syst. Zool.*, 21 : 364-374.
- NELSON G.J., 1994. Homology and systematics. In : B.K. Hall (ed.), *Homology : the hierarchical basis of comparative Biology*, p. 101-149, Academic Press, San Diego.
- NELSON G.J. & PLATNICK N., 1981. *Systematics and biogeography : cladistics and vicariance*. Columbia University Press, New York.
- PATTERSON C., 1982. Morphological characters and homology. In : K.A. Joysey & A.E. Friday (eds), *Problems of phylogenetic reconstruction*, Systematics Association special volume 21, p. 21-74, Academic Press, Londres et New York.
- PATTERSON C., 1988. The impact of evolutionary theories on systematics. In : D.L. Hawksworth (ed.), *Prospects in systematics*, Systematics Association special volume 36 : p. 59-91, Clarendon Press, Oxford.
- de PINNA M.C.C., 1991. Concepts and tests of homology in the cladistic paradigm. *Cladistics* 7 : 367-394.
- RIEPEL O., 2003. Popper and systematics. *Syst. Biol.*, 52 : 259-270.
- RIEPEL O., 2005. Le cohérentisme en systématique. In : P. Deleporte & G. Lecointre (coords.), *Biosystema* 24 – *Philosophie de la Systématique*, p. 115-126, Société française de systématique. Paris.
- TASSY P., 1991. *L'arbre à remonter le temps*. Christian Bourgeois éditeur, Paris.
- TASSY P., 1994. Les arbres phylogénétiques et l'ancêtre absent. In : P. Férida, D. Widlöcher & M. Wolf (eds.), *Les évolutions : phylogénèse de l'individu, colloque de la revue Internationale de Psychopathologie, bordeaux, 1993*, p. 99-110, Presses Universitaires de France, Paris.
- TASSY P. & BARRIEL V., 1995. L'homologie, l'arbre généalogique et le cladogramme : un apologue. *Bull. Soc. Zool. Fr.*, 120 : 361-378.
- TËMKIN I. & ELDREDGE N., 2007. Phylogenetics and material cultural evolution. *Current Anthropology*, 48 : 146-153.
- TORT P. (coord.), 1996. *Dictionnaire du darwinisme et de l'évolution* (3 volumes). Presses Universitaires de France, Paris.
- VAN WAEREBEKE D., GAULLIER V. & CHAIX R., 2008. *Espèces d'espèces*. Paris, Ex Nihilo, France 5, Arte France, CNRS Images.
- WAGNER W.H., 1961. Problems in the classification of ferns. In : D.L. Bailey (ed.), *Recent advances in botany — : from lectures and symposia presented to the IX International Botanical Congress Montreal, 1959*, volume 1, p. 841-844, University of Toronto Press, Montreal.
- WALLACE A R., 1856. Attempts at a Natural Arrangement of Birds. *Annals and Magazine of Natural History*, second series 105 : 193-216.

