

NOTION : FILIATION

Guillaume LECOINTRE¹

*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.)*

L'intuition selon laquelle les espèces sont généalogiquement affiliées entre elles diffuse dans la seconde moitié du 18^e siècle. On a coutume d'associer spontanément cette intuition au transformisme, mais sur ce point il faut rester prudent.

1 – Naissance de la filiation des espèces

Par exemple, chez le précurseur du transformisme qu'est Benoist de Maillet, 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 et 1751, mais ne restera jamais qu'implicite. Erasme 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

1. Systématicien, professeur au Muséum national d'histoire naturelle (UMR 7138 CNRS-UPMC-ENS-IRD-MNHN, « Systématique, adaptation, évolution ») et directeur du département « Systématique et évolution » du MNHN.

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 ?

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'Erasmus 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. Georges Buffon publia bien une carte généalogique des chiens en 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.

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'image de l'arbre apparaît en 1766 chez le zoologiste russe d'origine allemande Peter Simon Pallas et souligne les discontinuités de la nature. 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. Mais l'arbre implique-t-il chez ces auteurs une référence à la généalogie des espèces ? Chez Duschesne et Pallas, 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 chez Charles Darwin, en 1856 chez Alfred Russel Wallace. L'arbre qui s'impose, surtout dans le dernier tiers du 19^e siècle, est appelé « phylogénie » en 1866 par Ernst Haeckel.

La figure d'arbre la plus connue, peut-être juste après celle que Ernst Haeckel publia en 1874, est sans doute celle que publia Darwin dans *L'Origine des espèces* en 1859². Cette figure est à usages multiples et demande qu'on s'y arrête. 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.

1.1 – 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*

2. Reproduite et commentée dans Tassy, ce volume. (NdÉ.)

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, cf. Tort 1996 : 1189-1191). Voici comment Edmond Barbier traduit le passage mentionné ci-dessus :

[...] 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 hérité 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 la descendance desdites espèces ! 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 caractères³ 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 18^e siècle, ou bien ils jugent cette réponse insuffisante⁴. (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 actuel (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

3. **Sur la notion de caractère, en phylogénie contemporaine, cf. Barriel, ce volume.**
(NdÉ.)

4. **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.* »

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 soit le processus de cette transformation. Il suffit de dire «ascendance commune», l'adjectif «commune» jouant le rôle de «avec modification»⁵.

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:

5. Cf. aussi l'idée de réseau généalogique, cf. Samadi & Barberousse, ce volume. (NdÉ.)

[...] 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.)⁶

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 socio-politique 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 ».

1.2 – 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* et *divergente* ; • beaucoup d'espèces s'*éteignent* ; • généralement,

6. **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 ! (Cf. Tort 1996 : 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*.**

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 supraspécifiques 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 dessin, 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'apparement. 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 (cf. section 6) ; • 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 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).

1.3 – 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 distinctes⁷ 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 par un processus fait de hasard des variations⁸, hérédité⁹, sélection naturelle¹⁰ (le principe de divergence sera ici particulièrement important¹¹); 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

7. Cf. Lecointre, chapitre « Récit de l'histoire de la vie... », ce volume.

8. Cf. Heams (« Variation »), ce volume. (NdÉ.)

9. Cf. Heams (« Hérité ») et Huneman (« Sélection »), ce volume. (NdÉ.)

10. Cf. Huneman, ce volume. (NdÉ.)

11. Remarque : cependant, d'autres processus seraient à l'œuvre tout en maintenant une divergence, cela n'aurait pas d'impact sur les deux autres points ci-dessous.

acquis par voie d'ascendance commune ; il fournit le guide, le squelette qui charpentera le « cahier des charges » de la classification.

2 – L'homologie

2.1 – L'homologie : des paris sur la filiation

L'homologie est le discours sur les mêmes. L'homologie se propose en effet de donner le même nom à des attributs semblables portés par des êtres vivants distincts, qui plus est, des êtres vivants qui aujourd'hui ne se croisent pas entre eux. Il se pose alors deux questions fondamentales.

En vertu de quoi va-t-on reconnaître que deux attributs sont « les mêmes » ? En vertu de fonctions communes ? En vertu d'une même forme ? En vertu d'une même position dans l'organisme ? L'histoire du concept d'homologie nous montre que les trois paramètres ne se recouvrent pas (cf. tableau 1). Des attributs pris chez deux organismes différents peuvent avoir la même fonction sans se ressembler. Ils peuvent entretenir dans les organismes les mêmes relations de position sans pour autant avoir la même forme, avec ou sans la même fonction. Cela fait du concept d'homologie un concept dense.

		Structure			
		Sim+Phy	Sim+non Phy	Non Sim+Phy	Non Sim+non Phy
Fonction	Similaire	ANALOGIE HOMOGÉNIE MAG de mouton MAG de vache	ANALOGIE HOMOPLASIE Langue de fourmilier Langue de Pangolin	ANALOGIE Aile de chauve- souris Aile de pigeon	ANALOGIE Aile de mouche Aile de pigeon
	Différente	HOMOGÉNIE MAG de mouton Aile de pigeon	HOMOPLASIE MPG de mouton Aile de pigeon	HOMOGÉNIE Hyomandibulaire de carpe Étrier de vache	STRUCTURES NON COMPARABLES Aile de mouche Œil de pigeon

Tableau 1. Situations d'homologie entre deux organes, tenant compte du fait que la fonction est différente ou similaire (lignes), et du fait que la structure est similaire (« Sim ») ou non, héritée par ascendance commune (après confirmation par un arbre, « Phy ») ou non. MAG : membre antérieur gauche, MPG : membre postérieur gauche.

Selon quel objectif reconnaît-on deux structures comme les « mêmes » ? C'est une question qu'a oublié de se poser une certaine pédagogie de la « monstration »¹². Il suffirait de montrer à des élèves un squelette de membre antérieur de pigeon, un squelette de membre antérieur de chauve-souris et un squelette de membre antérieur d'homme pour que les élèves aient la révélation de l'évolution. On oublie que les champions de l'homologie qu'étaient Georges Cuvier et Richard Owen étaient fixistes. L'homologie reflétait des archétypes idéaux (cf. Hall 1994, Schmitt 2006) dont l'origine ne résidait aucunement

12. Cf. Fortin, ce volume. (NdÉ.)

dans des ancêtres communs. Le champion de l'homologie qu'était Étienne Geoffroy Saint-Hilaire y voyait, lui, sous un autre nom, finalement un moyen de rallier tardivement le transformisme. En somme, autant dans l'histoire des sciences que dans les classes de sciences naturelles d'aujourd'hui, ce n'est pas parce qu'on reconnaît des homologies d'organes partagées entre espèces distinctes que l'on pense immédiatement à l'évolution. Au contraire, c'est la correspondance des parties entre organismes qui a *besoin* d'être nourrie par un fond théorique pour être intelligible. Il se trouve que cette correspondance trouve aujourd'hui sa plus grande cohérence lorsqu'elle est nourrie de la théorie darwinienne de l'évolution.

Pour Darwin, il y avait des traits qui devaient rendre compte de l'ascendance commune entre plusieurs lignées, mais il y avait aussi des traits acquis en propre sur chacun des lignages qui risquaient de masquer les affiliations s'ils s'accumulaient trop. Sans méthode de construction d'arbres, impossible de démêler les autapomorphies, les simplésiomorphies, les synapomorphies, les homoplasies, et pour cause : les trois premières sont des concepts hennigiens, qui ne virent donc le jour qu'en 1950. Les méthodes classificatoires d'avant Hennig (1950, 1966) n'étaient ni formalisées, ni restreintes dans leur application à un échantillon fixe d'espèces dont on chercherait les relations d'apparentement. La recherche phylogénétique restait donc un exercice de synthèse hautement subjectif. Pour Hennig, la recherche des degrés relatifs d'apparentement ne peut se faire que sur un échantillon donné d'espèces. Dès lors, les êtres vivants se comportent comme des mosaïques de traits qui sont, pour certains, des signatures potentielles d'apparentement entre certains des membres de l'échantillon d'espèces à classer (et donc des marqueurs intéressants pour répondre à la question « Qui est plus apparenté à qui ? » : les futures « synapomorphies », états dérivés partagés), pour d'autres des traits trop généralement partagés pour la collection d'espèces que nous avons en mains (les futures simplésiomorphies), pour d'autres encore des traits trompeurs (les futures homoplasies) parce qu'ils auraient pu résulter d'évolutions convergentes (donc la ressemblance n'est pas ici héritée d'un ancêtre commun) ou parce que leur distribution ne résulte pas d'une acquisition ancestrale mais d'une disparition secondaire du trait dans une partie de l'échantillon. Pour assigner les traits ressemblants choisis à ces différentes catégories de traits, il va falloir construire un arbre dont la forme des branches va apporter les réponses aux questions. Mais comment obtenir un tel arbre ? C'est ce que nous verrons dans la section 3. Mais comment fonctionne l'homologie dans ce processus ?

Dans les enseignements de sciences naturelles du supérieur, on a longtemps appris qu'étaient homologues deux structures héritées par voie d'ascendance commune. On a longtemps appris également, et de façon séparée, qu'étaient homologues deux structures qui, prises chez des organismes

différents, entretenaient avec les structures voisines, les mêmes connexions, et ceci quelles que soient leurs formes et leurs fonctions (une définition qui remonte à Owen). Sans connecter l'une à l'autre, il était impossible de comprendre comment fonctionnait le concept. Selon Jonathan Wells (Lecointre 2004) et Michael Denton (Lecointre 1997), partisans du mouvement dit de l'«*Intelligent Design*», l'homologie est un concept circulaire et donc fallacieux : «*Pourquoi les manuels définissent-ils l'homologie comme une similarité due à une ascendance commune, puis déclarent que les homologies sont les preuves de l'ascendance commune, un argument circulaire déguisé comme une preuve scientifique ?*» Il y a ici, certes, un travestissement de la façon dont les scientifiques utilisent la notion d'homologie. Mais si ce travestissement fonctionne, c'est que le concept, dans son opérationnalité, est souvent mal enseigné et mal compris.

Une hypothèse d'homologie est d'abord un pari : partant de structures qui se ressemblent ou qui sont connectés aux structures voisines de la même façon (seconde définition ci-dessus, homologie primaire de de Pinna 1991), on fait initialement le pari qu'elles sont héritées d'un ancêtre commun. Mais ce pari, on peut le perdre ou le gagner. Lorsque l'on classe un échantillon d'espèces, on fait ce pari sur des dizaines, voire des centaines de caractères *en même temps*. Ces caractères sont inscrits dans un tableau qui comprend en ligne les taxons à classer et en colonne les caractères, les cases étant remplies de zéro et de un (ou autres symboles) en fonction de l'état dans lequel on trouve chaque caractère dans chaque taxon. L'exercice décisif, c'est la construction de l'arbre, c'est-à-dire, parmi l'ensemble des arbres possibles, le choix de celui qui maximise la contiguïté des états de caractères identiques. Cet arbre est alors le plus parcimonieux en hypothèses de transformation de caractères, car plus les états identiques sont rassemblés sur un tronc commun à plusieurs taxons qui le portent, moins il est besoin d'hypothèses d'acquisition de cet état le long des branches de l'arbre. Pris globalement, l'arbre choisi est celui qui maximise la cohérence entre ces multiples caractères. Et l'arbre le plus cohérent (le plus parcimonieux, le plus économe en hypothèses de transformation) va montrer que pour certains caractères, on a gagné le pari, tandis que pour d'autres, on l'a perdu. Dans le premier cas, l'homologie sera confirmée et l'on parlera d'homologie secondaire (de Pinna 1991) ou d'homogénéité (Lankester 1870). Ces homologies deviendront alors des arguments en faveur de l'apparemment exclusif des espèces qui les portent (ce sont les fameuses synapomorphies).

Par exemple, dans l'arbre phylogénétique le plus parcimonieux obtenu avec une collection de trente espèces de vertébrés parmi lesquels un échantillon de quatre oiseaux (canard, poulet, colibri et albatros) et de multiples caractères, le bréchet, l'un de ces caractères, est acquis une seule fois sur la branche qui réunit ces derniers : le bréchet est une structure homologue par filiation chez

le canard, le poulet, le colibri et l'albatros et cette homologie (cette synapomorphie) constitue un argument en faveur de leur apparentement exclusif. Dans le second cas, l'homologie sera infirmée, on parlera alors d'homoplasie (ressemblance non acquise par ascendance commune). Reprenons l'exemple précédent. La collection d'animaux comprenait également deux espèces de chauve-souris, un chat et un chien. Le même arbre le plus parcimonieux montre que la présence d'une oreille avec pavillon, de poils, de mamelles et d'une mandibule constituée d'un seul os (le dentaire) n'est acquise qu'une seule fois sur la branche qui réunit les chauves-souris avec le chat et le chien. Pour tous ces caractères, le pari d'homologie est gagné (ce sont des synapomorphies). En revanche, on constate que, dans cet arbre, le caractère « membre antérieur réalisant une aile » n'est pas acquis une seule fois, mais deux fois : une fois sur la branche menant aux quatre oiseaux, une autre fois sur la branche menant aux chauve-souris. Si l'on avait fait le pari que le caractère « membre antérieur réalisant une aile » était homologue chez les oiseaux et les chauves-souris, il se trouve à présent qu'on l'a perdu. Il en est de même pour le bec. Dans cet échantillon, il y avait également une tortue, située dans l'arbre sur une branche qui n'est pas celle des oiseaux. Le bec des tortues et le bec des oiseaux, que l'on a pu supposer initialement homologues, ne le sont finalement pas : ils ont acquis plusieurs fois indépendamment au cours de l'évolution. Insistons sur le fait que c'est la parcimonie globale ayant guidé le choix de l'arbre qui est responsable de toutes ces conclusions. C'est en confondant l'homologie comme pari et l'homologie comme résultat – et donc en occultant la dimension du pari – que Wells et Denton ont fait de l'homologie un concept d'usage circulaire. Mais dans la mesure où l'on peut perdre le pari, il n'y a pas circularité. Les concepts d'homologie primaire et secondaire de Mário de Pinna (1991) sont donc utiles en cela qu'ils réinscrivent l'homologie dans son opérationnalité en tant que concept participant à la construction d'un arbre phylogénétique.

2.2 – La notion de plan : les homologies antiphylogénétiques

La notion de plan s'enracine dans les sciences naturelles dès lors que des hommes trouvent des points communs aux diverses espèces, depuis Aristote et Galien jusqu'à Geoffroy Saint Hilaire en passant par Léonard de Vinci, Newton, Diderot, Daubenton, Buffon. Beaucoup de ces auteurs – Diderot ou Buffon par exemple – parlent davantage de « prototype » que de « plan ». Cette recherche de l'unité anatomique va tout particulièrement baigner la seconde moitié du 18^e siècle et la première moitié du 19^e (cf. Schmitt 2006). En France, elle va trouver son apogée avec Étienne Geoffroy Saint Hilaire, qui recherche une unité de plan pour tout le vivant, et Georges Cuvier, qui utilise aussi la notion de plan d'organisation, mais à l'échelle des grands embranchements seulement : Cuvier combat l'idée que cette notion puisse être étendue à l'ensemble du vivant (la polémique entre les deux hommes fit grand bruit

en 1830). En Allemagne, au sein de la *Naturphilosophie*, Johann Wolfgang von Goethe et ses successeurs utilisent la notion de «type» dans le sens de «type d'organisation», abstraction méthodologique permettant une lecture des organismes s'affranchissant des multiples variations qui brouillent ce qu'ils ont de commun. Ils sont suivis en cela par les Anglais Joseph Henry Green, Robert Edmond Grant, Robert Knox et Richard Owen, avec qui un point culminant sera atteint – Owen recherche les «archétypes». Les types permettent de reconnaître une certaine unité de structure au sein d'espèces variées, mais leur signification varie d'un auteur à l'autre. Pour ne prendre que deux exemples, les variations autour d'un thème unique ne relèvent pour Goethe que de nécessités fonctionnelles liées aux habitudes qu'ont prises les êtres vivants dans les divers milieux qu'ils habitent; il n'y pas de transformisme. En revanche, pour Owen ou pour Karl Ernst von Baer, il y a possibilité de transformation des espèces si elle est limitée à l'intérieur d'un type, même si ces auteurs n'admettaient pas la sélection naturelle de Darwin. Si la notion de plan a été féconde dans l'histoire de l'anatomie comparée, on peut se demander à quoi elle sert encore à l'heure de l'analyse phylogénétique. Elle s'est encore épanouie encore durant les cent ans qui ont suivi la publication de *L'Origine des espèces*, à une époque où l'on classait avant de rechercher les «affinités» (nous dirions aujourd'hui les «relations d'apparementement», même si elles n'étaient pas alors distinguées des relations d'ancêtre à descendant): en effet, les embranchements et les classes étaient définis par des archétypes idéaux et non pas des synapomorphies, ce, même dans un contexte évolutionniste. Mais après l'avènement de la construction d'arbres phylogénétiques, la classification devint postérieure à l'élucidation des degrés d'apparementement. Les embranchements, les classes et autres groupes n'eurent plus besoin d'être fondés sur une représentation vitrifiée d'un vivant idéal, qu'il soit plan ou type. La hiérarchie dans la répartition des attributs est rendue par l'arbre phylogénétique, lequel est l'expression même *à la fois* d'un emboîtement de taxons et d'une évolution en mosaïque des caractères, ce qui rend caduque la nécessité de cristalliser une série d'attributs en un plan.

La question fondamentale du «plan d'organisation» est de savoir si, aujourd'hui, ce dernier sert à l'investigation scientifique ou seulement à la pédagogie, constituant en quelque sorte un moyen mnémotechnique pour les apprentis naturalistes. Le problème est qu'utiliser la notion de plan dans un cours de zoologie ou de botanique, c'est cristalliser dans les mémoires ce qui semble implicitement s'être cristallisé dans la nature. Or le plan ne fait que figer ensemble en un *organisme idéal* un certain nombre d'attributs comme caractéristiques d'un groupe donné. En cela, il mélange ce qui relève de la reconnaissance («voici par quels attributs on reconnaît un vertébré ou un annélide» au-delà des multiples variations qui en brouillent la perception) et

ce qui relève de la pensée évolutive, en assignant au plan abstrait l'image de ce qu'a dû être l'ancêtre hypothétique du groupe.

Dans l'histoire des sciences, on est passé allègrement de l'un à l'autre, du critère pratique de reconnaissance à l'inférence évolutive. Or, nous l'avons dit, la phylogénie nous enseigne rétrospectivement que les organismes ont subi une évolution mosaïque et que certains attributs du plan qui est dans nos têtes se « désolidarisent » facilement des autres attributs. Par exemple, on peut facilement élaborer un « plan » de vertébré. On y trouvera yeux, vertèbres, mandibule, nageoire caudale. Mais ce plan échouera à servir de critère de reconnaissance pour les *Astyanax* qui n'ont pas d'yeux. On pourra alors objecter la mauvaise foi qu'il y aurait à ne pas vouloir tenir compte du fait que les *Astyanax* ont tout de même tous les autres attributs du plan. Il suffit d'émettre l'hypothèse *ad hoc* d'une perte secondaire des yeux. Le problème est que cette hypothèse n'est plus du tout de même nature. Il ne s'agit plus de constater et de reconnaître mais *d'inférer*. Le plan bascule donc forcément dans le discours évolutionniste. Il en ira de même pour les *Mola* et tous les tétrapodes qui n'ont pas de nageoire caudale. Qu'à cela ne tienne, on peut alors élaborer un plan tétrapode, sans nageoire caudale mais avec quatre membres marcheurs. Le problème n'est que déplacé : bien des tétrapodes ne sont pourvus que des résidus de pattes (*Bipes*) ou de ceintures (pythons, boas), voire totalement dépourvus de pattes (amphisbènes, gymnophiones, orvets, serpents). Sans parler des oiseaux qui n'ont pas quatre membres marcheurs, mais seulement deux : certains de leurs membres n'ont rien de marcheurs. Dans ce dernier exemple, on émettra l'hypothèse *ad hoc* que ces membres se sont transformés.

On voit donc qu'en faisant incursion dans le discours évolutionniste, le « plan » sort de sa vocation pratique d'outil mnémotechnique pour imposer des pertes secondaires ou des modifications ultérieures et s'autodéclarer alors forcément « plan ancestral », lequel est une mauvaise méthode d'inférence d'un ancêtre. Cette méthode se passe de phylogénie car elle est idéaliste et non formalisée, elle est construite non pas sur la maximisation formelle de la cohérence d'observations (comme on le ferait pour l'inférence phylogénétique), mais à coup d'hypothèses *ad hoc*. La bonne méthode pour inférer un ancêtre doit être transparente et cohérente : elle consiste tout simplement à retenir les états de caractères inférés au nœud de l'arbre phylogénétique où se trouve l'ancêtre auquel on s'intéresse.

Par ailleurs, rien n'indique que la combinaison d'attributs inscrite dans le « plan » ait existé, surtout parce que la notion de plan interdit à l'ancêtre qui est représenté de façon tacite d'avoir eu des traits dérivés propres. En concentrant en un seul organisme ce qui est pensé arbitrairement comme « primitif », l'organisme abstrait n'a pas « droit » à ses états de caractères dérivés spéciaux.

En d'autres termes, si l'ornithorynque était le plan mammalien, il n'aurait pas droit à son bec, ni à ses palmures, ni à ses éperons venimeux.

Enfin, la métaphore du plan renvoie à une ambiguïté du mot lui-même : le plan est aussi l'expression d'une intention finalisée ou une programmation ; on parle de « cinquième plan », de « plan divin », de « plan B », etc. La fixation idéale de traits de caractères présentés comme « constants », « généraux » ou « primitifs » en un organisme virtuel pourrait être avantageusement qualifiée de « carte » en remplacement de la métaphore du plan. En effet, il s'agit avant tout de montrer une combinaison d'attributs et leurs connexions entre eux. La carte ne porte pas en elle-même d'idée d'intention à l'origine de l'être qui est montré, elle est une simple représentation schématique et utilitaire de ce qui est. Cependant, même dans ce cadre-là, si la carte prétend schématiser ce qui est, il faudrait alors qu'elle représente un organisme concret en le nommant, et non une combinaison idéale de caractères.

Parce qu'elle fige la nature dans des étapes imaginaires là où, partout, il n'y a que transformations continues et réversibles, la notion de plan est impropre à une approche phylogénétique du vivant : un organisme n'est pas le reflet d'un plan idéal, mais une mosaïque unique de caractères qui est le fruit de la contingence de son histoire évolutive.

3 – La construction de l'arbre : Willi Hennig

Entre 1859 et 1950, on sait dessiner des arbres mais on ne sait pas construire d'arbre. On dispose bien, depuis Darwin (1859), du niveau théorique (celui de la généalogie présumée du vivant) mais pas du niveau pratique (comment fait-on ?). Willi Hennig, entomologiste allemand, publie sa systématique phylogénétique en 1950 en allemand, mais ne sera entendu de la communauté internationale qu'en 1966 lors de la traduction de son livre en anglais. Quatre principes fondamentaux caractérisent la pensée hennigienne : (i) on ne classe jamais qu'un échantillon précis d'espèces ; (ii) seuls les traits dérivés, innovants au sein de l'échantillon, peuvent conduire à un regroupement taxonomique reflétant l'apparentement ; (iii) pour savoir si un trait est dérivé, il faut polariser le caractère concerné, c'est-à-dire aller voir dans quel état il est dans l'extragroupe (et faire cela pour tous les caractères disponibles) ; (iv) les groupes paraphylétiques ou polyphylétiques ne sont que des artefacts classificatoires fondés sur des traits primitifs partagés ; et une taxonomie phylogénétique ne doit admettre que les groupes monophylétiques (clades).

3.1 – Construction des caractères

En systématique, on pratique la biologie comparée qui consiste à détecter la ressemblance entre les structures présentées parmi un échantillon d'organismes, en les comparant entre eux. On utilise une grille de lecture théorique commune à tous les organismes de notre échantillon, qui va permettre de

détecter les ressemblances et les différences entre organismes à travers le principe des connexions. Selon une définition vieille de cent cinquante ans, *sont homologues des structures qui, prises chez des organismes différents, entretiennent avec les structures voisines les mêmes connexions et les mêmes relations topologiques, et ceci quelles que soient leur forme et leur fonction.*

Prenons l'exemple classique du membre antérieur des tétrapodes. Le squelette du membre antérieur du dauphin, de la chauve-souris et de l'homme ne sont pas identiques, ne fonctionnent pas de la même manière, mais il est pourtant possible de détecter un agencement commun de leurs différentes pièces constituantes. Le radius du dauphin est homologue du radius de la chauve-souris, en dépit de leurs formes et tailles respectives très dissemblables : dans les deux cas ils sont connectés à un segment proximal unique lui-même connecté à la ceinture scapulaire (l'humérus). Ils sont aussi connectés à un second os parallèle (ulna), puis à des pièces carpiennes plus distales. Si, de plus, de par la forme du radius, nous pouvons mettre en évidence que le radius de la chauve-souris ressemble plus à celui de l'homme qu'à celui du dauphin, *on peut émettre une hypothèse d'homologie* (ou homologie primaire), c'est-à-dire qu'on fait *le pari qu'à cette ressemblance anatomique correspond une communauté d'ascendance. C'est l'arbre phylogénétique qui, une fois reconstruit, donnera la réponse au pari.* Il dira si les structures supposées homologues sont bel et bien héritées d'un ancêtre commun hypothétique (homologie secondaire, ou homologie confirmée), ou bien si elles sont apparues plusieurs fois indépendamment. L'homologie primaire est une hypothèse d'homologie à partir des connexions. L'homologie secondaire, c'est l'homologie par ascendance commune, confirmée par l'arbre.

L'ontogenèse nous révèle parfois également l'origine des organes. Il existe un parallèle entre le développement embryonnaire et le déroulement phylogénétique, remarqué depuis le 19^e siècle. Si l'ontogenèse récapitule la phylogenèse, à une communauté d'origine embryologique doit correspondre, en première approximation, une communauté d'origine phylogénétique. L'embryologie peut donc également conduire à formuler des hypothèses d'homologie primaire. Par exemple, une mandibule de lézard est constituée de plusieurs os, notamment le dentaire, l'articulaire, l'angulaire ; cette mandibule s'articulant au crâne au niveau de l'os carré. Or, chez un mammifère, la mandibule est formée du seul os dentaire, lequel s'articule au crâne au niveau d'un autre os, le squamosal. L'articulaire, l'angulaire et l'os carré sont passés dans l'oreille. En effet, deux lignes d'argumentation se corroborent très fortement pour établir des homologies entre des os de la mandibule reptilienne et des os de l'oreille mammalienne. L'une est paléontologique, l'autre est embryologique. Toute une série de fossiles du Permien et du Trias montrent une position de plus en plus postérieure de l'articulaire et de l'angulaire, jusqu'à leur intégration à

la sphère auditive tandis que l'articulation de la mandibule au crâne devient dento-squamosale. L'embryologie des marsupiaux montre le détachement de pièces cartilagineuses de la mandibule vers la sphère auditive. Ces données établissent que l'angulaire de la mâchoire d'un lézard est l'homologue de l'os tympanique des mammifères, que l'articulaire du lézard est l'homologue au marteau, l'un des trois osselets de l'oreille moyenne mammalienne conduisant les vibrations du tympan vers la fenêtre ovale, et que l'os carré du lézard est l'homologue de l'enclume, le second des trois osselets de l'oreille moyenne des mammifères. Sans l'embryologie ni les fossiles, il eût été fort difficile d'établir une hypothèse d'homologie entre des os mandibulaires et des os de l'oreille moyenne.

3.2 – On ne classe jamais qu'un échantillon

De nos jours, la classification naturelle unique que réalisent les chercheurs est la classification phylogénétique du vivant. Pour comprendre Hennig, il faut commencer par réaliser qu'on se propose de classer un échantillon du vivant, et jamais tout le vivant d'un coup. Cette remarque semble anodine mais elle est capitale pour garantir transparence et reproductibilité des procédures, ainsi qu'un véritable contrôle du principe d'économie d'hypothèses. Cet échantillon, on le veut «représentatif» d'un problème posé. La valeur et la signification des attributs des organismes (les caractères) peuvent ne pas être les mêmes d'un échantillon à l'autre. Par exemple, le fait d'avoir quatre membres ne sert à aucun regroupement au sein des mammifères, tandis qu'au sein des vertébrés, la présence des quatre membres est une innovation évolutive qui signe l'apparementement des tétrapodes.

3.3 – Les innovations évolutives de l'échantillon sont les marqueurs de l'apparementement exclusif

On se propose donc toujours de retrouver les relations de parenté au sein d'un nombre limité d'espèces, un échantillon, par exemple un lézard, un serpent, un crocodile, un poulet, un canard et une pie. Cet échantillon possède toujours, dans l'arbre phylogénétique de la vie, un ancêtre commun à tous les représentants choisis. Ceci reste vrai même si on ne connaît pas l'ordre de leurs branches : il suffit de redescendre jusqu'à ce qu'on parvienne à la première branche commune. Appelons cet ancêtre abstrait X. Il s'agit d'enregistrer des caractères, ou attributs de ces organismes, pour lesquels on peut distinguer au moins deux états au sein de l'échantillon. Ces états sont, en première approximation, supposés homologues (cf. section précédente). Par exemple, le bourgeon épidermique donne de l'écaille chez certains de ces organismes, et de la plume chez d'autres. Écaille et plume sont homologues *a priori* en tant que dérivés de l'épiderme dont les bourgeons se ressemblent.

Une classification «classique» conduirait, pour ce seul caractère intitulé «bourgeon épidermique», à créer un groupe lézard + serpent + crocodile d'un

côté, avec comme argument le partage des écailles, et de l'autre un groupe poulet + canard + pie fondé sur le partage de la plume. Pour Hennig, ce n'est pas la bonne manière de procéder. Si nous émettons une hypothèse d'homologie entre écaille et plume, et si nous assumons pleinement l'évolutionnisme dans la manière de construire une classification, alors *l'un des deux états est dérivé de l'autre*. Pour l'échantillon d'espèces, cela signifie que l'un des deux états a été acquis depuis X (l'état dérivé) et l'autre a été acquis dans l'arbre avant X (l'état primitif). *Celui des deux états acquis avant X ne peut pas servir à faire des regroupements au sein de notre échantillon, puisqu'il est déjà présent en dehors, sur d'autres branches de l'arbre de la vie*. Si nous faisons un groupe sur le partage d'un tel état, notre groupe serait incomplet, oubliant tous ceux qui sont à l'extérieur. Par contre, *il faut faire des regroupements sur la présence commune d'états de caractères dérivés*. Ils constituent des innovations évolutives dans l'échantillon, et elles seules peuvent fournir des apparentements exclusifs et donc des groupes complets.

3.4 – Polariser les caractères par l'ontogénie

Déterminer l'état primitif, cela se nomme polariser le(s) caractère(s). Comment sait-on lequel des deux états est dérivé de l'autre? Il existe deux critères utilisés actuellement. Le critère ontogénétique et le critère de l'extra-groupe.

Le critère ontogénétique utilise le fait que l'ontogenèse récapitule la phylogenèse – disons, en première approximation. Le développement embryonnaire évoluant dans le temps phylogénétique par additions successives de séquences nouvelles de développement, les états de caractères d'apparition précoce dans le développement sont distribués plus généralement dans le vivant que ne le sont les états de caractères d'apparition plus tardive. Par exemple, dans le déroulement phylogénétique, nous avons été vertébrés (il y a 500 millions d'années [m.a.]) avant d'être tétrapodes (il y a 380 m.a.), avant d'être des primates hominoïdes (il y a 25 m.a.). Dans le temps embryologique, la colonne vertébrale apparaît avant les membres, qui apparaissent eux-mêmes avant la régression de la queue qui caractérise les primates hominoïdes (cf. Lecointre & Le Guyader 2006). Celui des deux états qui donne naissance à l'autre est donc l'état primitif, celui qui en découle est l'état dérivé. Le bourgeon épidermique préfigure déjà l'écaille mais se transforme en plume. On peut même, expérimentalement, transformer les écailles des tarses des poulets en plumes en ayant injecté de l'acide rétinoïque dans l'embryon à un certain stade de son développement. Tout semble donc indiquer que la plume est une écaille dérivée. Dans ce cas, le groupe que l'on va constituer est celui des espèces de l'échantillon possédant l'état dérivé, à savoir la plume, donc le groupe poulet + canard + pie.

3.5 – Polariser les caractères par comparaison avec l'extra-groupe

Le critère extra-groupe consiste à prendre un référentiel, à aller voir ce qui se passe à l'extérieur de l'échantillon que l'on se propose de classer. Aucune construction scientifique ne se fait sans postulats. Une classification n'échappe pas à cette règle, mais il faut que les postulats soient explicites. On énonce alors clairement le postulat suivant : la tortue est extérieure à l'échantillon lézard + serpent + crocodile + poulet + canard + pie. Ce postulat fait de la tortue un extra-groupe. Observons à présent l'épiderme de la tortue. On y trouve de l'écaille. C'est donc que l'écaille était, dans l'arbre de la vie, déjà présente avant X. L'écaille est l'état primitif, la plume est l'état dérivé. Par conséquent, nous allons créer un groupe au sein de l'échantillon, qui signera l'apparement exclusif poulet + canard + pie, «exclusif» signifiant à l'exclusion du lézard, du serpent et du crocodile.

3.6 – Assumer ses postulats

Il convient de faire trois remarques au sujet de l'extra-groupe. La première, c'est qu'il est généralement recommandé d'utiliser plusieurs extra-groupes de manière à bien argumenter la polarisation des caractères. La seconde, c'est que si nous nous étions trompés d'extra-groupe en utilisant, par exemple, une autruche, nous aurions posé la plume comme état primitif. Le groupe créé aurait été lézard + serpent + crocodile. Cela signifie que la procédure en elle-même ne gère pas la pertinence des postulats. *Comme partout en sciences, la pertinence des postulats ne relève que de la fiabilité des connaissances extérieures à l'expérience, lesquelles doivent être assumées par le chercheur.* Si l'un des postulats est faux, le résultat a peu de chances d'être vrai, mais alors ce n'est pas la procédure qui est en cause. Enfin, il est clair qu'on ne fonde pas une classification sur un seul caractère, mais qu'on en utilise *plusieurs simultanément*. Cela nous conduit à l'exercice suivant.

3.7 – La procédure standard

Après l'exposé des principes hennigiens, il convient à présent d'exposer la procédure élémentaire qui permet aujourd'hui, à l'aide de l'informatisation des procédures, de construire une classification. Dans un souci de simplicité et de place, nous proposons de classer un nombre minimal d'espèces. Rappelons-nous que la question fondamentale est «Qui est plus proche de qui?». À deux espèces, cela n'a pas de sens. On peut répondre à partir de trois espèces. Avec trois espèces, deux d'entre les trois peuvent être plus proches entre elles que chacune ne l'est de la troisième. Enfin, si nous tenons compte de la section précédente, il faut un référent extérieur pour polariser les caractères : l'extra-groupe. Le minimum d'espèces à considérer est donc de quatre. Prenons par exemple une chauve-souris, un homme, un pigeon. L'ancêtre commun à tous dans l'arbre de la vie (l'ancêtre X ci-dessus) est l'ancêtre commun hypothétique à tous les amniotes. L'extra-groupe sera donc pris en dehors des amniotes, par

exemple une truite. La question est donc : parmi la chauve-souris, l'homme et le pigeon, quels sont les deux les plus apparentés ? En laissant toujours la truite à l'extérieur, ce qui fait partie des postulats, il y a trois réponses possibles, exprimables ci-dessous sous forme de parenthèses, mais aussi sous forme d'arbres (fig. 2) :

((chauve-souris, homme) pigeon) : b
 ((chauve-souris, pigeon) homme) : c
 ((pigeon, homme), chauve-souris) : d

La procédure moderne de construction d'une classification s'emploie donc à explorer un champ d'hypothèses possibles.

3.8 – Coder les caractères

Il faut à présent explorer des caractères pour évaluer l'argumentation de chacune d'elles. Les observations faites sur les différents états des caractères vont être codés à l'aide de chiffres et inscrites dans un tableau appelé matrice de caractères¹³ (fig. 2a). Par convention d'écriture, l'état trouvé dans l'extra-groupe est noté «0». Attention, le zéro ne signifie pas une absence de caractères, mais simplement l'état dans lequel il est dans l'extra-groupe. Nous pouvons proposer :

1. Sacs aériens : sous forme de vessie natatoire (0) ; sous forme de poumons alvéolés fonctionnels (1)
2. Appendices pairs : sous forme de nageoires rayonnées (0) ; sous forme de membres marcheurs (1)
3. Constitution de la mandibule : plusieurs os (0) ; un seul os, le dentaire (1)
4. Glandes mammaires : absence (0) ; présence (1)
5. Ailes : absence (0) ; présence (1)
6. Dents : présence (0) ; absence (1)
7. Colonne vertébrale : présence (0) ; absence (1)

Compte tenu de la question posée ci-dessus, on constate d'emblée que certains caractères ne servent à rien. Le fait de posséder une colonne vertébrale (caractère 7) ne discrimine nullement deux d'entre les trois. C'est la même chose pour les caractères 1 et 2. La chauve-souris, l'homme et le pigeon possèdent tous les trois les caractères 1, 2 et 7 sous le même état. On dit que ces caractères ne sont pas informatifs au regard de la question posée. Ceci valait d'être souligné, parce que c'est de cette manière que les chercheurs sélectionnent leurs caractères. Dans le gisement de caractères que constitue une collection d'organismes, ils ne prélèvent que les caractères potentiellement informatifs. Cependant, parmi les caractères 3, 4, 5 et 6, il en reste un qui ne

13. Cf. Barriel, ce volume. (NdÉ.)

pourra servir à établir un regroupement. C'est le caractère 6. En effet, il semble discriminant puisqu'il conduit à une ressemblance commune à l'homme et à la chauve-souris qui tous deux possèdent des dents. Mais sachant que la présence de dents est déjà observée dans l'extra-groupe, on en déduit qu'il s'agit ici un état primitif partagé et l'absence de dents un état dérivé. On se souvient de la règle de Hennig qui consiste à ne pas établir de groupes sur la base de caractères primitifs partagés mais seulement sur la base d'un état dérivé partagé. L'état dérivé étant l'absence de dents, le caractère 6 ne fait donc que regrouper le pigeon avec lui-même. En fait, du point de vue algorithmique, un caractère informatif est celui qui présente au moins deux états (ce qui n'est pas le cas pour le 7), chacun étant représenté au moins deux fois (ce qui n'est pas le cas des caractères 1, 2, 6). Les seuls caractères informatifs sont donc les caractères 3, 4 et 5.

3.9 – Placer les transformations des caractères sur les arbres

On constate que les caractères se contredisent entre eux. En effet, les caractères 3 et 4 montrent les états dérivés communs à l'homme et à la chauve-souris, et le caractère 5 montre un état dérivé commun à la chauve-souris et au pigeon. Comment s'en sortir ? En observant deux règles : la première, *en plaçant dans chacun des arbres possibles les transformations des caractères* (fig. 2 b, c, d), la seconde, en respectant *le principe d'économie d'hypothèses*. Par exemple (fig. 2, arbre b), on va tenter de placer l'endroit dans l'arbre où le caractère 3 change d'état, c'est-à-dire passe d'une mandibule à plusieurs os (état « 0 » trouvé chez la truite) à une mandibule constituée du seul os dentaire (état « 1 »). Ce passage deviendra une *hypothèse de transformation* (symbolisé par un petit cercle de couleur, fig. 2 b, c, d). Cette action doit respecter le principe d'économie d'hypothèses, selon lequel on ne va pas multiplier les transformations au delà du simple nécessaire. Par exemple, dans l'arbre b, on pourrait vouloir faire apparaître la mandibule constituée du seul os dentaire une fois sur la branche menant à l'homme et une autre fois sur la branche propre à la chauve-souris. Cela respecterait les observations. Mais on peut faire « moins cher » : en plaçant plus simplement cet événement au segment de branche commun à l'homme et à la chauve-souris, réalisant ainsi une seule transformation au lieu de deux. En réalisant cet exercice sur tous les caractères, on observera que pour certains caractères il n'est pas possible de faire moins de deux transformations. Dans l'arbre b, nous sommes forcés de faire apparaître les ailes deux fois, car il n'y a pas de branche menant exclusivement au couple chauve-souris + pigeon : une fois sur la branche propre à la chauve-souris et une fois sur la branche propre au pigeon. Il faut répéter cet exercice sur chacun des trois arbres possibles, de telle manière qu'à la fin de cet exercice, tous les caractères auront été interprétés en termes de transformations évolutives dans chacun des arbres possibles.

	1. Sacs aériens	2. Appendices pairs	3. Mandibule	4. Glandes	5. Ailes	6. Dents	7. Colonne vertébrale
Truite	0	0	0	0	0	0	0
	1	1	1	1	1	0	0
Homme	1	1	1	1	0	0	0
Pigeon	1	1	0	0	1	1	0

Figure 2a. Matrice de caractères. Attention : 0 ne veut pas nécessairement dire « absence » ; il ne s'agit que d'une convention de codage consistant à indiquer « comme dans l'extra-groupe », l'extra-groupe étant ici la truite.

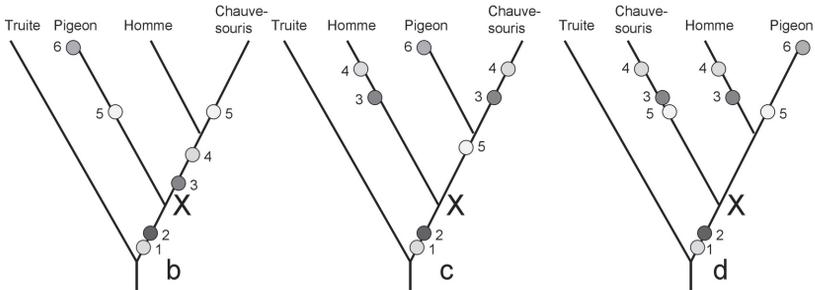


Figure 2b, c, d. Les trois arbres possibles montrant toutes les relations d'apparement possibles entre trois taxons et un extra-groupe. X : ancêtre hypothétique aux trois membres du groupe d'étude. Chaque boule est une hypothèse de transformation de l'état « 0 » à l'état « 1 » pour le caractère considéré. L'arbre b coûte 7 hypothèses de transformation, l'arbre c 8 hypothèses et l'arbre d 9 hypothèses. On retient l'arbre b.

3.10 – Appliquer le principe de parcimonie

Le principe d'économie d'hypothèses, utilisé dans toutes les sciences, va conduire à retenir l'arbre b, parce qu'il ne « coûte » globalement que 7 hypothèses de transformations, tandis que les deux autres arbres possibles sont moins parcimonieux (8 hypothèses pour c et 9 hypothèses pour d). Bien entendu, les chercheurs travaillent sur beaucoup plus d'espèces et plus de caractères. Ils utilisent alors des algorithmes qui réalisent la procédure exposée ici, à l'aide d'ordinateurs. Ceux-ci sont en fait indispensables, car le nombre d'arbres possibles à explorer augmente drastiquement avec le nombre d'espèces à classer.

3.11 – Deux groupes-frères ont le même rang

L'arbre retenu est celui qui est le plus parcimonieux. Il fournit une classification, et en même temps les réponses aux paris sur les homologies. Dans l'arbre b, on dit que l'homme et la chauve-souris sont groupes-frères. *Deux groupes-frères ont le même rang.* Ainsi, dans l'arbre phylogénétique des vertébrés, les chondrichthyens (vertébrés cartilagineux : requins, raies) et les ostéichthyens (vertébrés osseux) sont groupes-frères : alors chondrichthyens

et ostéichthyens sont de même rang (par convention, on décidera qu'il s'agit de deux classes).

3.12 – Nommer les clades

Clade est synonyme de « groupe monophylétique ». Homme et chauve-souris sont tous deux inclus dans un clade. *On peut donner un nom à chaque clade*. Dans l'exemple simplifié qui est le nôtre, ce nom pourrait être « Mammifères ». Pigeon, homme et chauve-souris sont inclus dans un clade plus inclusif, comprenant le premier, et donc de rang supérieur. S'il fallait le nommer, il s'agirait ici des Amniotes. En fait, dans tout arbre phylogénétique, chaque segment de branche interne correspond à un groupe nommé qui englobe tout ce qui est en aval de cette branche (Lecointre & Le Guyader 2001). Si nous sommes capables de construire des phylogénies et de suivre cette règle, tous les groupes ainsi créés seront monophylétiques, et la classification qui en découlera sera réellement phylogénétique, comme l'avait souhaité Darwin. Dans le cahier des charges des classifications actuelles, l'arbre phylogénétique guide donc, voire fournit la classification.

3.13 – L'arbre retenu nous enseigne quels attributs sont hérités d'un ancêtre commun exclusif

Il y a trois types de ressemblance. *La première* est celle due à un caractère primitif partagé, trop général pour pouvoir opérer une classification au sein de l'échantillon présent. Par exemple, le fait d'avoir des vertèbres font se ressembler nos trois espèces, mais cela ne nous dit rien sur « qui est plus proche de qui ». De même, la présence de dents ne regroupe aucunement la chauve-souris et l'homme à l'exclusion du pigeon, puisque les dents sont présentes dans l'extra-groupe. *La deuxième* est due à des caractères dérivés partagés. Ces ressemblances sont celles qui permettent de créer des classifications phylogénétiques. Nous avons fait le pari (« homologie primaire ») que la mandibule faite du seul os dentaire et les mamelles avaient été acquises d'un ancêtre commun. L'arbre le plus parcimonieux nous enseigne que nous avons eu raison. Ces ressemblances-là sont des homologies confirmées, ou « homologies secondaires ». *La troisième* est la ressemblance qui correspond à des caractères dérivés non obtenus par ascendance commune. En effet, au cours du déroulement évolutif, des caractères qui se ressemblent beaucoup peuvent être acquis plusieurs fois indépendamment dans des lignées différentes. Par exemple, les mammifères « ont fait de la taupe » trois fois : une fois en Australie (la taupe marsupiale), une fois en Eurasie (les taupes placentaires qui sont des Eulipotyphles) et une fois en Afrique australe (les Chrysochloridés) sans que les ancêtres respectifs de ces animaux n'aient eu une forme de taupe. Ces convergences évolutives sont dues au fait qu'une bonne partie des contraintes physiques que subissent les espèces sont les mêmes partout. Ainsi, les ailes de notre échantillon ne sont pas homologues en

tant qu'ailes : on a appris à voler une fois chez les chauves-souris et une autre fois chez les oiseaux. Ces phénomènes expliquent pourquoi, dans une matrice de caractères, les caractères se contredisent les uns les autres ; nous avons fait des paris sur l'homologie des caractères, mais pour certains d'entre eux nous nous sommes trompés, sans pouvoir détecter notre erreur *a priori*. Cela se produit toujours ; il n'existe pas de matrice de caractères qui ne contienne pas de la contradiction interne. *Cette contradiction entre caractères, on l'appelle homoplasie*, ou incohérence de la matrice. Un arbre donné possède une certaine quantité d'homoplasies, ou ressemblances non héritées d'un ancêtre commun. Ici, l'arbre b présente une homoplasie, l'aile.

L'analyse cladistique, le plus souvent suivie d'une classification phylogénétique, permet de prendre conscience qu'*une classification n'est plus un art qui enferme son praticien dans l'autorité de son expertise, mais un démarche scientifique qui requiert des postulats, des observations formalisées, transparentes à tout autre classificateur ou utilisateur, des règles du jeu parmi lesquelles l'économie d'hypothèses, des décisions explicites, et, surtout, l'exploration d'un champ des possibles* (les différents arbres possibles). En somme, la classification phylogénétique, par l'explicitation qu'elle force, est capable de produire de la connaissance objective, ce qui est le propre des connaissances scientifiques.

4 – 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.
- 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. En vertu du « *descent with modification* », le « qui partage quoi avec qui » exprime finalement des degrés relatifs d'apparementement entre individus concrets. 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 et métaphorique. Quelle est la nature de cette métaphore ? Il s'agit d'une grande généalogie théorique, constituée de milliards d'ancêtres et de descendants reliés entre eux par des relations génétiques et à jamais inconnaissables 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, 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 et phylogénie (tableau 2).

ARBRE GÉNÉALOGIQUE DE LA VIE	ARBRE PHYLOGÉNÉTIQUE
Prospectif : se parcourt dans le sens des processus de l'évolution (« process »)	Rétrospectif : reconstitué à partir du partage des structures (« patterns »). Se parcourt à rebours.
Ancêtres inconnaissables individuellement	Ancêtres partiellement reconstitués, mais restant des portraits-robots
Montre « qui descend de qui » entre individus abstraits	* Montre « qui est apparenté à qui » entre individus concrets * Montre « qui descend de qui » entre ancêtres reconstitués et descendants concrets
Statut théorique	Statut heuristique
Relève de l'ontologie	Relève de l'épistémologie
Tableau 2. Arbre généalogique de la vie et arbre phylogénétique : deux arbres à ne pas confondre.	

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 de 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.

5 – Qu'est-ce que la phylogénie ?

La définition de la phylogénie comporte deux 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.

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 est trompeur (Dayrat 2003). 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 créa 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 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, va reprendre le mot pour parler de la grande généalogie des êtres organisés (Darlu & Tassy 1993, Dayrat 2003, 2005). Aucune de ces deux définitions ne sont celles qui sont retenues aujourd'hui. Car 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 manquait à Darwin (cf. Dupuis 1978, 1986, 1992).

Aujourd'hui, 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. Cependant, ce point a été âprement discuté et les interprétations divergent (cf. Brower 2000, Rieppel 2005, Lecointre, ce volume).

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 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. 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 symplésiomorphies, 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, qui vient du fait que les taxons sont définis par des caractères homologues (Patterson 1982, 1988, Nelson 1994), est cruciale pour le systématique (le plus souvent morphologiste) dont le travail est de créer des taxons monophylétiques, donc d'identifier des homologies secondaires. Elle semble cependant superflue pour le généticien, 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 homologues 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 : ils leur reprochent de ne pas le faire véritablement. 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 moléculaires¹⁴ » 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 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 préférence, parce que ces mêmes préférences de processus prétendument connus *a priori* s'accompagnait, dans les pratiques passées de la systématique éclectique, qu'un manque criant de formalisation. Cependant,

14. Cf. la section 7 de Barriel, ce volume. (NdÉ.)

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 (Rieppel 2003), certainement plus abductives qu'hypothético-déductives¹⁵. Le « cohérentisme » des cladistes peut alors se muer en « fondhérentisme » (en anglais, « *foundherentist* » le terme est de Haack 2000; cf. Rieppel 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 (Rieppel 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 elle est théoriquement chargée (Rieppel 2005). Parmi ces connaissances sont incluses celles sur les processus biologiques. Si le systématicien était conscient de *tous* ses postulats et qu'il devait en exclure

15. Cf. Lecointre (« *Récit de l'histoire de la vie...* »), ce volume.

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 fondhéntiste consciente est préférable à une pureté cladiste (version 1980) épistémologiquement illusoire et empiriquement impossible à tenir.

Enfin, l'alternative est étonnante lorsque des systématiciens parlent de « découvrir des processus ». Les systématiciens 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 partir d'elle... Et elle est rarement fiable dans son entièreté à partir d'un seul jeu de données (tout systématicien 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'œuvre¹⁶. 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 systématiciens ! 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, 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 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

16. Cf. Lecointre (« *Récit de l'histoire de la vie...* »), ce volume.

du «*descent with modification*». C'est la première raison pour laquelle 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 (cf. section 8). Elle n'est pas faite pour cela, tout simplement. La seconde raison en est que 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. 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 homologies. 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 apparentement), 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'apparentement. 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.

6 – La forme de l'arbre de la vie

Depuis que des transferts « horizontaux » de matériel génétique¹⁷ ont été prouvés entre lignages bactériens en principe généalogiquement séparés depuis

17. Cf. la section 5.1 de Heams (« Hérité », ce volume. (NdÉ.)

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* » tendrait alors à devenir un réseau réticulé, au moins dans certaines de ses parties. 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 portent¹⁸. Le problème se pose avec d'autant plus d'acuité les reconstructions phylogénétiques impliquent de reconstituer des divergences anciennes. Nous avons vu plus haut qu'il convenait 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 apparements 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).

18. Cf. Samadi & Barberousse, ce volume. (NdÉ.)

7 – La filiation, cahier des charges de toute classification biologique

Lamarck ne donne pas de programme sérieux à l'exercice classificatoire. En 1809, la classification doit rester avant tout opérationnelle, un « produit de l'art » qui ne « tient réellement rien de la nature », et en 1820 elle doit refléter « l'ordre le plus naturel » qui n'est pas spécifié.

7.1 – La contrainte de monophylie est-elle incluse chez Darwin (1859) ?

Celui qui fixera réellement – c'est-à-dire explicitement et durablement – le cahier des charges des classifications sera Darwin (cf. citation, section 1.1). 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 porte en germe une autre révolution qui ne portera ses fruits qu'au 20^e 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, programmée (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. 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 gommeaient les relations d'apparentement.

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 chirodien) ; 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'apparement 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 (cf. le début de ce chapitre), 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'apparement. 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 relient 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 (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 importants¹⁹. » 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 prédarwiniens, 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 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^e siècle comme

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

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 *utilise 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 écrivit 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 une lignée 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 à la lignée 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, 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 maintient. 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 20^e 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 19^e 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 Ernst Mayr en 1974 alors qu'il défendait les grades. La distinction négative reptile/oiseaux masque l'apparement 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 considère qu'un groupe taxonomique évolue et donne naissance à un autre, parce que l'essence du reptile se maintient 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 tétons. 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 19^e siècle. Scalisme, anthropocentrisme, maintien de la tradition et du progrès, les 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.

8 – 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 inconnus 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 taxinomique. 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 taxinomiques, 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 la 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 *abstrait*s 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*. Dayrat (2005) répond que non, mais avec deux défauts : il ne distingue pas dans son article les ancêtres abstraits 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 pré-tentions poppériennes assez naïves, mais une autre épistémologie peut être proposée pour la cladistique (comme déjà signalé) 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) ne va pas au fond du problème. Les méthodes qu'il préconise n'identifient pas 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) considérerait 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, 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 infraspécifique 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 *abstrait*s restent *indispensables théoriquement*.

Références bibliographiques

- BROWER A. (2000), "Evolution is not a necessary assumption of cladistics", *Cladistics*, 16: 143-154.
- DARLU P. & TASSY P. (1993), *Reconstruction phylogénétique. Concepts et méthodes*, Paris, Masson. Disponible sur <http://sfs.snv.jussieu.fr/publications_sfs.shtml>.
- DARWIN C. (1859), *The Origin of Species, First Edition*, London, John Murray. Rééd. Penguin Classics, London, Penguin Books, 1995.
- DARWIN C. (1860), *The Origin of Species, Second Edition*, London, John Murray. Rééd. Oxford World's Classics, Oxford, Oxford UP, 1998.
- DARWIN C. (1872), *L'origine des espèces*, Paris, C. Reinwald, 1880 (traduit par E. Barbier sur l'édition définitive – 6^e édition – de *The Origin of Species*, London, Murray, 1872).
- DAYRAT B. (2003), "The roots of phylogeny: How did Haeckel build his trees?", *Syst. Biol.*, 52(4): 515-527.
- DAYRAT B. (2005), "Ancestor-descendant relationships and the reconstruction of the tree of life", *Paleobiology*, 31(3): 347-353.
- DE BARBANÇOIS-VILLEGONGIS C.H. (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.
- DE PINNA M.C.C. (1991), "Concepts and tests of homology in the cladistic paradigm", *Cladistics*, 7: 367-394.
- DELEPORTE P. & LECOINTRE G. (2005), «Philosophie de la systématique», *Biosystema* 24, Paris, Société française de systématique.
- DUPUIS C. (1978), «Permanence et actualité de la systématique: la "systématique phylogénétique" de W. Hennig (Historique, discussion, choix de références)», *Cahier des naturalistes*, *Bull. N. P. n. s.* 34(1) : 1-69.
- DUPUIS C. (1986), «Darwin et les taxonomies d'aujourd'hui», in P. Tassy (coord.), *L'ordre et la diversité du vivant*, Paris, Fayard, Fondation Diderot : 215-240.
- DUPUIS C. (1988), «Le taxinomiste face aux catégories», *Cahier des naturalistes*, *Bull. N. P. n. s.* 44(3) : 49-109.
- DUPUIS C. (1992), «Regards épistémologiques sur la taxinomie cladiste. Adresse à la XI^e session de la Willi Hennig Society (Paris, 1992)», *Cahier des naturalistes*, *Bull. N. P. n. s.* 48(2) : 29-56.
- DURIS P. & GOHAU G. (1997), *Histoire des sciences de la vie*, Paris, Nathan université.
- ELDRIDGE N. & CRACRAFT J. (1980), *Phylogenetic patterns and the evolutionary process. Method and theory in comparative Biology*, New York, Columbia UP.
- ESCOBAR-PARAMO P., SABBAGH A., DARLU P., PRADILLON O., VAURY C., DENAMUR E. & LECOINTRE G. (2004), "Decreasing the effects of horizontal gene transfer on bacterial phylogeny: the *Escherichia coli* case study", *Mol. Phylogenet. 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, New York, Columbia UP : 1-36.
- GAYON J. (1992), *Darwin et l'après-Darwin*, Paris, Kimé.
- 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*, Malden, Mass., Blackwell : 226-236.
- HAECKEL E. (1866), *Generelle Morphologie der Organismen*, Berlin, Reimer.
- HAECKEL E. (1874), *Anthropogenie*, Leipzig, Engelmann.
- HALL B.K. (ed.) (1994), *Homology, the hierarchical basis of comparative biology*, New York, Academic Press.
- HENNIG W. (1950), *Grundzüge einer Theorie der phylogenetischen Systematik*, Berlin, Deutscher Zentralverlag.

- HENNIG W. (1966), *Phylogenetic Systematics*, Urbana, University of Illinois Press.
- HUSON D.H. & BRYANT D. (2006), "Application of Phylogenetic Networks in Evolutionary Studies", *Mol. Biol. Evol.*, 23(2) : 254-267.
- KIRIAKOFF S. (1963), «Les fondements philosophiques de la systématique biologique», in La classification dans les sciences, Bruxelles, Duculot : 61-88.
- LAMARCK J.-B. (1994), *Philosophie zoologique [1809]*, Paris, Flammarion.
- LAURIN M. (2008), *Systématique, paléontologie et biologie évolutive moderne : l'exemple de la sortie des eaux chez les vertébrés*, Paris, Ellipses.
- LANKESTER E.R. (1870), "On the use of the term homology in modern zoology, and the distinction between homogenetic and homoplastic agreements", *Ann. Mag. Nat. Hist.*, (4)6 : 34-43.
- LECLERC M.C., V. BARRIEL, G. LECOINTRE, B. DE REVIERS. (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. (1997), «Evolution et molécules : Denton en crise», in P. Tort (dir.), *Pour Darwin*, Paris, PUF : 693-729.
- LECOINTRE G. (2004), «Jonathan Wells, Phillip Johnson et Michael Denton : les formes instruites de l'anti-science», in J. Dubessy, G. Lecointre & M. Silberstein (dir.), *Les matérialismes (et leurs détracteurs)*, Paris, Syllepse : 511-543.
- LECOINTRE G. & DELEPORTE P. (2005), "Total evidence requires exclusion of phylogenetically misleading data", *Zoologica Scripta*, 34(1) : 101-117.
- LECOINTRE G. & LE GUYADER H. (2006), *Classification phylogénétique du vivant*, 3^e éd., Paris, Belin.
- MAYR E. (2004), *Après Darwin*, Paris, Dunod.
- NELSON G.J. (1972), "Comments on Hennig's 'phylogenetic systematics' and its influence on ichthyology", *Syst. Zool.*, 21(4) : 364-374.
- NELSON G.J. (1994), "Homology and systematics", in B.K. Hall (ed.), *Homology: the hierarchical basis of comparative biology*, San Diego, Academic Press : 101-149.
- NELSON G.J. & PLATNICK N. (1981), *Systematics and biogeography: cladistics and vicariance*, New York, Colombia UP.
- PATTERSON C. (1982), "Morphological characters and homology", in K.A. Joysey & A.E. Friday (eds.), *Problems of phylogenetic reconstruction*, London and New York, Academic Press : 21-74.
- PATTERSON C. (1988), "The impact of evolutionary theories on systematics", in D.L. Hawksworth (ed.), *Prospects in Systematics*, Syst. Assoc., Spec. Vol. 36 : 59-91.
- RIEPEL O. (2003), "Popper and systematics", *Syst. Biol.*, 52 : 259-270.
- RIEPEL O. (2005), «Le cohérentisme en systématique», in P. Deleporte & G. Lecointre (coord.), *Philosophie de la systématique*, Biosystema 24, Paris, Société française de systématique : 115-126.
- SCHMITT S. (2006), *Aux origines de la biologie moderne : l'anatomie comparée d'aristote à la théorie de l'évolution*, Paris, Belin.
- TASSY P. (1991), *L'arbre à remonter le temps*, Paris, Christian Bourgeois éditeur.
- TASSY P. (1994), "Les arbres phylogénétiques et l'ancêtre absent", in P. Férida & D. Widlöcher (dir.), *Colloque de la Revue internationale de psychopathologie*, Paris, PUF : 99-110.
- TASSY P. & BARRIEL V. (1995), «L'homologie, l'arbre généalogique et le cladogramme : un apologue», *Bull. Soc. Zool. Fr.*, 120(4) : 361-378.
- TORT P. (dir.) (1996), *Dictionnaire du darwinisme et de l'évolution*, Paris, PUF.
- WAGNER W.H. (1961), "Problems in the classification of ferns", in *Recent Advances In Botany*, University of Toronto Press : 841-844.

