

PASCAL TASSY 1991

L'arbre à remonter le temps : les rencontres de la systématique et de l'évolution

Complété par

PIERRE DARLU, PASCAL TASSY, P. 1993 (3ème édition).

La reconstruction phylogénétique du vivant : concepts et méthode. Paris : Bélin. Version électronique 2004

2011, mars 2019

L'analyse cladistique devient un instrument essentiel de la reconstruction des dynamiques sociétales. Il convient donc de remonter aux sources

Je propose ici, sur la base des textes de Pascal Tassy (fig. 1), une relecture «anthropologique» des fondements théoriques de la cladistique. En tant que procédure d'ordination structurale, ce type d'analyse taxonomique à vocation historique, utilisée en paléontologie, me semble en effet un instrument performant pour une approche de l'histoire des sociétés humaines sans pour autant se mouler dans une conception évolutionniste dogmatique. La cladistique permet en effet d'affiner et d'enrichir l'approche classificatoire des sociétés sur des bases réfutables.

La thèse que je défends ici est que la cladistique peut être adaptée à l'étude de l'histoire des phénomènes culturels humains, indépendamment de tout présupposé évolutionniste. Cette ouverture méthodologique nous a été suggérée par Valentine Roux lors d'une conférence donnée le 6 décembre 2011 au Cercle genevois d'archéologie sur l'évolution de la technique du colombin au tour au Proche Orient.

Je développerai cette réflexion afin de prolonger et approfondir plus particulièrement les tentatives récentes de classification et de catégorisation de divers types de sociétés au plan général (Testart 2005a ; Gallay 2006) ou plus particulièrement africaniste (Gallay 2011a).

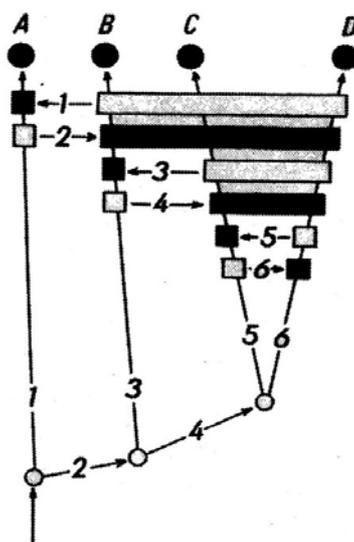


Fig. 1. Schéma d'argumentation phylogénétique de Willi Hennig (tiré de *Phylogenetic Systematics* 1966). Relations de parenté entre quatre taxons (A à D) fondées sur 6 caractères.

En noir : état dérivé (apomorphe) des caractères; en blanc : état primitif (plésiomorphe) Tassy 1991, fig. 15.

Taxonomie et phylogénie en biologie : un historique

Tout recours au domaine culturel implique des classifications des caractéristiques retenues dont il convient d'approfondir les fondements épistémologiques. Pour comprendre les enjeux de cette question un détour par l'histoire des classifications en biologie est nécessaire, ces problématiques ayant des champs d'application beaucoup plus larges.

Ce rapide historique des relations en taxonomie et histoire des espèces vivantes en biologie évolutive est nécessaire pour situer les enjeux de la cladistique.

Le dualisme entre « somme des modifications » (la taxonomie des êtres vivants) et indicateurs des liens de descendance (la phylogénie) n'a cessé de hanter les systématiciens. Les contradictions entre approches systématiques concurrentes au cours du XX^e siècle naîtront toutes de la façon dont est abordée la similitude. S'agit-il de mesurer les quantités de différence et de similitude ou de repérer les ressemblances qui témoignent de la filiation ? Depuis *L'Origine des espèces* de Darwin, pendant près d'un siècle, les pratiques phylogénétiques mêleront les deux approches sans jamais bien distinguer l'une de l'autre (voir également Gallay 1986).

Nous pouvons distinguer trois phases dans l'évolution des conceptions des relations entre taxonomie et phylogénie.

Années 40-50 : le triomphe de l'évolutionnisme néodarwinien ; années 60 : la contestation phénétique ; années 70 : « la révolution cladistique ». En réalité, la réflexion cladistique remonte jusqu'à la fin des années 30, lorsque l'entomologiste allemand Willi Hennig essaie de résoudre, après Darwin, le problème de la coordination des catégories de la classification. C'est en 1950 que Hennig publie, en langue allemande, dans son maître livre : *Grundzüge einer Theorie der phylogenetische Systematik*, les principes de la théorie de la systématique phylogénétique.

A la suite de Pascal Tassy on attribuera par convention l'étiquette « évolutionniste » à ce qui a trait à la systématique néo-darwinienne, puisque c'est ainsi qu'elle s'est appelée elle-même ; et l'étiquette « phylogénétique » sensu stricto à ce qui a trait à la systématique cladistique, puisque c'est sa dénomination originale. Le terme phylogénétique reste portant polysémique car il est souvent utilisé au sens large pour désigner toute préoccupation portant sur l'enchaînement historique des espèces vivantes.

Phase 1. Les classifications néo-darwiniennes

Le dualisme entre « somme des modifications » (la taxonomie des êtres vivants) et indicateurs des liens de descendance (la phylogénie) n'a cessé de hanter les systématiciens. Les contradictions entre approches systématiques concurrentes au cours du XX^e s. naîtront toutes de la façon dont est abordée la question des similitudes. On peut distinguer quatre étapes jalonnant ces réflexions.

Dans le néodarwinisme la recherche de la structure de la diversité biologique est subordonnée à celle des mécanismes censés générer la variabilité et, plus encore, induite par cette dernière. La recette taxonomique néodarwinienne des constructions d'arbres est fondée sur un outil permettant le classement. L'homologie (cf. infra) est la clé de l'interprétation de la similitude

et des produits, les taxons, associant des caractères, mais également des théories sous-jacentes sur les mécanismes : adaptation, sélection, gradualisme, équilibres ponctués, etc. (Gould 2002).

Phase 2. Les classifications phénétiques

Critiqué dans les années 60 par les phénéticiens adeptes de la taxinomie numérique et contesté de tous côtés, l'arbre néodarwinien est aujourd'hui en voie de disparition.

La systématique évolutionniste a été attaquée en raison de sa composante phylogénétique et son interprétation historique. Les taxons reconnus par les néodarwiniens ne pouvaient être construits que sur la base d'hypothèses évolutionnistes vagues confondant grades et clades. Ces griefs posés, les phénéticiens ont proposé de substituer la recherche de relations fondées sur la similitude globale à celle de toute autre relation, notamment évolutionniste. Les caractères devaient désormais être les plus nombreux possibles. La somme des modifications, c'est-à-dire la ressemblance globale, dûment calculée au moyen de procédures mathématiques appropriées et souvent sophistiquées pouvait seule, dans cette optique, donner des résultats taxonomiques fiables et stables. Les phénéticiens ne pouvaient admettre qu'un seul caractère (le trait apomorphe du cladisme) valait mieux qu'un indice résumant une centaine de mesures.

En fait le but original de la taxonomie numérique n'était pas de proposer une nouvelle méthode d'analyse évolutionniste, mais seulement de nouvelles règles de classification. En cela, nombre de phénéticiens n'ont jamais prétendu que le phénogramme était un arbre phylogénétique et ont laissé aux phylogénéticiens la construction d'une généalogie évolutionniste comprise comme l'identification des séquences de branchements de l'arbre.

Phase 3. Les classifications cladistiques

Darwin lui-même, et nombre de ses successeurs néodarwiniens, ont utilisé la somme des modifications comme outil taxonomique de grande efficacité. Dans l'*Origine*, Darwin précise :

"Je crois que l'arrangement des groupes dans chaque classe, d'après leurs relations et leurs de gré de subordination mutuelle doit, pour être naturel, être rigoureusement généalogique".

A la fin des années 60 les paléontologues avancent que la sophistication d'un appareil mathématique ne suffit pas à valider une taxonomie représentée sous la forme d'un arbre. D'un point de vue évolutionniste, l'approche phénétique est désormais considérée comme une impasse. La simple similitude graphique (des paires de taxons réunis par des lignes qui se rejoignent en des points de branchement hiérarchiques) facilite abusivement le transfert conceptuel du phénogramme en arbre phylogénétique (Gallay 1986, p. 69). Mais le phénogramme n'a pas de racine. Il ne raconte pas une histoire avec une origine et une suite d'évènements. Il illustre seulement des proximités morphologiques. A l'inverse des arbres néodarwiniens, lointains hérités de l'idéologie victorienne du progrès, le phénogramme est un schéma dichotomique où n'apparaît évidemment aucune relation d'ancêtre à descendant, mais uniquement des proximités entre paires de taxons.

Subordonner des considérations phylogénétiques à l'établissement de groupements fondés sur la similitude globale implique en effet que la parenté phylogénétique est un décalque de la distance morphologique, que les « points de descendance » sont identiques à la somme des modifications », c'est-à-dire que la vitesse d'évolution est la même pour tous les caractères des organismes étudiés, une position qui est naturellement intenable sur le plan scientifique.

La systématique phylogénétique de Hennig devient alors célèbre sous le nom de « cladisme », un néologisme conçu en 1963 par Ernst Mayr pour qualifier, quelque peu ironiquement, l'étude des branchements généalogiques.

Le point de départ de l'entreprise cladistique est essentiellement pragmatique : comment simplifier l'histoire extrêmement complexe qu'est l'histoire de la descendance en rejetant la notion de grade comme stade évolutif et en fournissant un cadre scientifique testable aux constructions taxonomiques ?

Cladistique

La cladistique (du grec ancien *κλάδος* *klados*, signifiant « branche »), aussi appelée systématique phylogénétique, est une théorie de classification phylogénétique. Dans son sens premier relevant des sciences biologiques, la cladistique est l'étude de la classification des êtres vivants, selon leurs relations de parenté, dans un cadre évolutionniste. Elle repose sur la construction de groupes monophylétiques dits clades qui incluent un ancêtre commun et l'ensemble de sa descendance.

L'approche cladistique se distingue radicalement de la taxonomie numérique : les caractères sont pris individuellement et n'évoluent pas de la même façon, à la même vitesse. Dans un même taxon, certains caractères sont dérivés, d'autres sont primitifs. Ces derniers ne possèdent pas, en bloc, des caractères tous restés primitifs ou tous devenus évolués. L'homme a un cerveau évolué, mais des membres à cinq doigts restés primitifs.

L'analyse cladistique repose sur trois postulats :

1. Les caractères évoluent indépendamment,
2. leurs vitesses d'évolution ne sont pas constantes,
3. leur évolution n'est pas irréversible.

Aucun de ces postulats n'implique que le cas inverse est impossible, ce qui ne révèle aucun modèle évolutif contraignant.

Dans notre perspective anthropologique la cladistique apporte un progrès considérable face aux simples taxonomies empiriques des phénomènes culturels, et ceci sur plusieurs plans :

- Une taxonomie cladistiques est un modèle ou une structure, soit, dans notre nomenclature, une régularité relevant du constructivisme et de l'empirisme logique (Gallay 2011b).
- En tant que classement, elle incorpore des données de première importance sur le plan historique en tenant compte de certains liens entre caractères « primitifs » et caractères « dérivés » introduisant une dimension dynamique supplémentaire dans les taxonomies.
- Les bases du classement, choix et définition des caractères, liens entre caractères primitifs et dérivés, etc. peuvent être discutés, modifiés et/ou réfutés sur la base de données factuelles. Elles peuvent donc tenir compte des multiples débats anthropologiques portant sur la diversité culturelle.
- Les liens entre caractères primitifs et dérivés peuvent être expliqués par des mécanismes et des processus relevant des dynamiques sociales et culturelles et donc d'une anthropologie dynamique dont Georges Balandier (1974) s'est fait le promoteur pour ce qui touche l'Afrique.
- La structure dégagée ne répond à aucun dogmatisme ou a priori évolutionniste et relève de l'empirisme. Elle est susceptible de s'actualiser dans divers scénarios locaux relevant de l'histoire.

- Ces diverses caractéristiques relèvent d'une épistémologie générale située en dehors du domaine des sciences biologiques. Elles sont donc applicables à n'importe quel phénomène se modifiant cours de l'histoire et ne préjuge d'aucune équivalence ou identité postulée entre espèces biologiques et phénomènes culturels tels que sociétés, systèmes techniques, systèmes sociaux, systèmes politiques, systèmes religieux, etc. La situation est donc identique à celle de la taxonomie numérique qui peut être appliquée à toutes sortes de phénomènes naturels ou culturels.
- Les diverses caractéristiques tentent de cerner la globalité des sociétés. La liste retenue peut évidemment être discutée et modifiée selon le jugement porté sur le caractère stratégique de chacun des critères face à l'histoire. Il n'y a pas un seul classement possible, mais une infinité. Chaque arborescence possible illustre un point de vue particulier porté sur ce qui est jugé important pour l'évolution des sociétés (fig. 2 et 3).

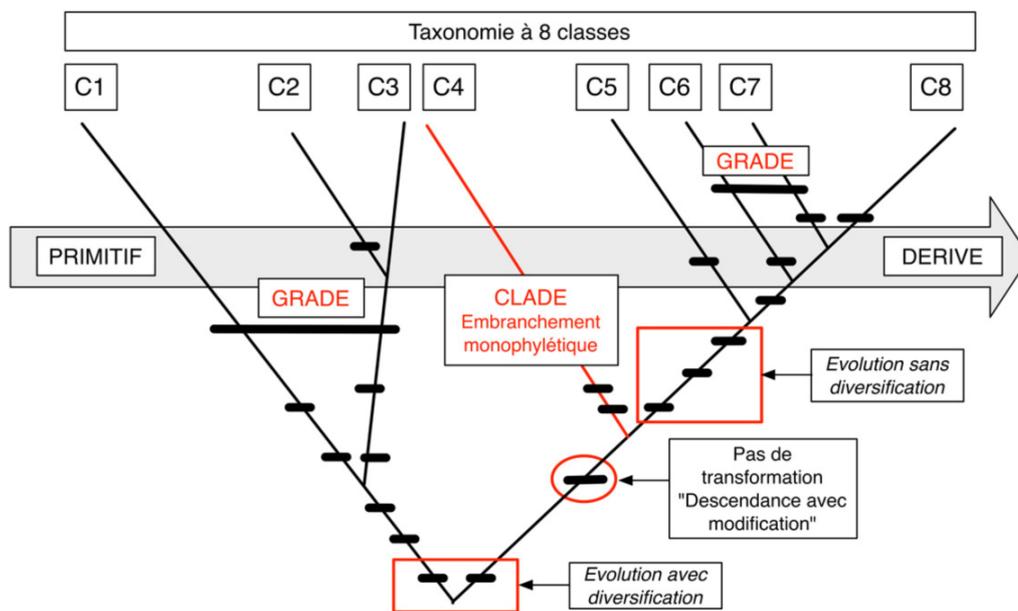


Fig. 2. Fondements théoriques de l'analyse cladistique dans un système à huit classes (C1 à C8). Les classes sont définies par des pas de transformation résultant de processus de descendance avec modification et définissant des embranchements monophylétiques. A l'opposé les grades caractérisent des modifications partagées par plusieurs clades.

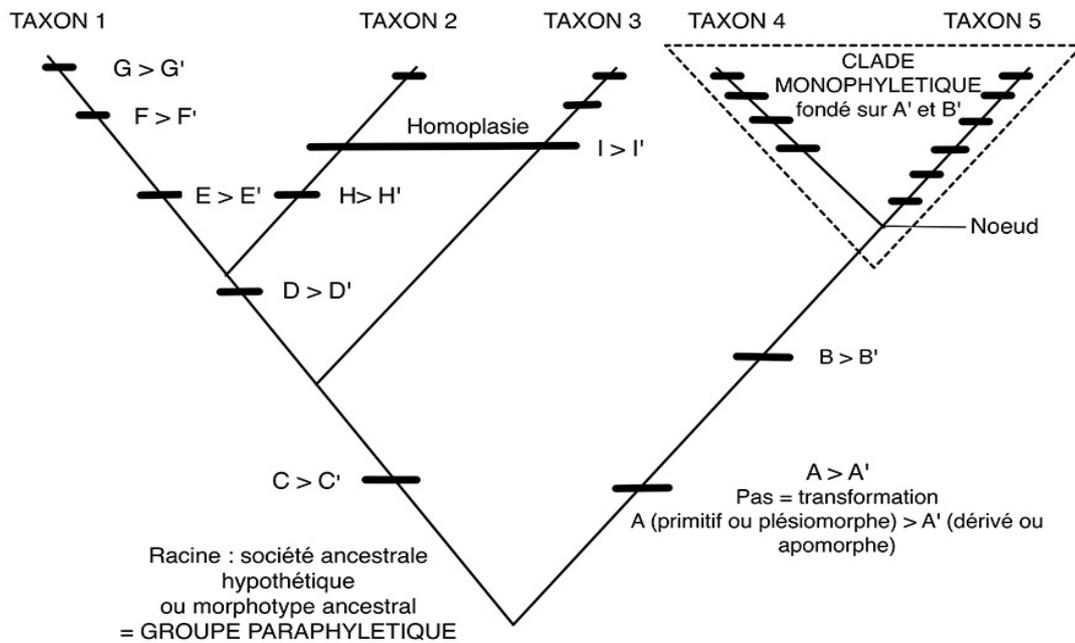


Fig. 3. Terminologie cladistique.

Caractères

Les critères, attributs ou traits des classifications seront tels caractères ou tel ensembles de caractères, choisis à l'issue d'une démarche raisonnée, pour leur puissance de regroupement.

L'association dans un même taxon de traits primitifs et de traits dérivés explique que la similitude globale ne saurait être, au mieux, qu'une approximation de la parenté. Sur ce point le cladisme s'oppose fondamentalement à la taxonomie numérique. L'histoire des taxons ne se reconstitue pas à partir des taxons eux-mêmes mais à partir de leurs caractères.

Pendant longtemps le point de vue dominant du néodarwinisme avait été que l'on ne pouvait étudier utilement des caractères que si l'on connaissait les processus les ayant mis en place. Une telle exigence a comme conséquence ultime de nier toute légitimité à l'entreprise de la systématique. La prééminence à la fois logique et historique du cladogramme est fondée sur le fait qu'il n'est pas nécessaire de proposer des hypothèses sur les mécanismes ou les processus pour construire une arborescence et que, pour un seul cladogramme, plusieurs scénarios sont possibles, inscrits dans les documents historiques.

Des combinaisons parcimonieuses de critères peuvent se révéler difficilement explicables, d'un point de vue strictement anthropologique et/ou historique. Ce constat doit alors motiver un retour aux caractères et la présentation d'un nouveau « codage » permettant des agencements différents.

Caractères primitifs et dérivés

Toute reconstruction phylogénétique tient de l'hypothèse. Hypothèse sur la transformation des caractères, hypothèse sur les groupements phylogénétiques. Mais la nature de ces hypothèses réside dans leur testabilité. Elles sont ouvertes à la réfutation : ce sont des hypothèses scientifiques.

La cladistique se distingue des autres domaines de la systématique car elle cherche à établir les relations de parenté strictes entre les taxons en distinguant des caractères primitifs (plésiomorphies, proches de l'état primitif) et des caractères dérivés (apomorphies, éloignés de l'état primitif). Un groupe monophylétique ne se fondera que sur le partage de caractères dérivés propres, également appelés synapomorphies.

La notion de polarisation et de sens de transformation est ici importante. Un caractère pouvant exister sous deux états: primitif ou dérivé, il faut pouvoir déterminer quel état est apomorphe et lequel est plésiomorphe. Pour cela on peut utiliser plusieurs critères : mobilisation d'une référence ou d'un groupe externe, données d'ordre historique, congruence, etc. Le critère ontologique impliquant une homologie entre développement ontogénétique et évolution des espèces, propre au domaine biologique, n'intervient pas ici.

Soit deux états a et b d'un même caractère ; on ne connaît pas le sens de transformation entre a et b on a donc $a \leftrightarrow b$. On peut donc admettre deux sens de transformation : $a \rightarrow b$ ou $b \rightarrow a$. Dans le cas $a \rightarrow b$, a se transforme en b, donc a est plésiomorphe par rapport à b et seul b nous renseigne sur les relations de parenté. Dans le cas $b \rightarrow a$, b se transforme en a, b est donc plésiomorphe par rapport à a et seul a nous renseigne sur les relations de parenté.

Les états primitif et dérivé sont, rappelons-le, des notions relatives. Un caractère est dérivé à tel nœud du cladogramme et devient primitif pour les nœuds subordonnés.

Caractères dérivés sériés

Un même caractère peut présenter un enchaînement de caractères dérivés. Les relations entre les états multiples d'un caractère peuvent être de plusieurs types. Elles peuvent être non ordonnées ou, au contraire, être ordonnées. Dans ce dernier cas on parle de « série de transformations » du caractère. Cette série peut être linéaire (au sens de sans bifurcation) ou non linéaire (avec bifurcation). Les caractères dont les états sont ordonnés seront également appelés additifs.

Dans des *relations non ordonnées* chaque état peut se transformer directement en tout autre état, chaque transformation ne comptant que pour un pas. Les relations ordonnées ou additives correspondent à des séries de transformations linéaires ou non linéaires, orientées ou non orientées : - série linéaire non orientée :

$$\mathbf{a - b - c - d}$$

Un caractère à états multiples est dit linéaire (donc également additif) quand on peut passer successivement d'un état à un autre. Chaque transformation valant un pas, cela implique nécessairement que le passage d'un état extrême (ici a) à l'autre (ici d) demande autant de pas qu'il y a d'états moins un : le passage de l'état a à l'état d (ou l'inverse) demande ici 3 pas.

- *Série linéaire orientée (ou dirigée) :*

$$\mathbf{a > b > c > d}$$

Dans cette série linéaire additive b est apomorphe par rapport à a et plésiomorphe par rapport à c. Il convient de ne pas faire de confusion entre caractères ordonnés (concernant l'ordre des transformations) et caractères orientés (concernant le sens des transformations).

- *Série non linéaire non orientée:*

La série non linéaire présente les relations entre états multiples sous forme d'un arbre. Toutes les transformations ne comptent pas toutes pour un même nombre de pas. Toutes ces

représentations des transformations, ou codages, ne sont pas « neutres », puisque l'on code une hypothèse phylogénétique qui est immédiatement apparente grâce à la forme arborescente du graphe de transformations des états du caractère.

Matrices taxons /caractères

En général les possessions de caractères sont résumées dans un tableau appelé matrice taxons/caractères donnant pour chaque caractère de chaque taxon son état. Par convention, pour chaque caractère "x", l'état primitif est noté x et l'état dérivé x' (ça aurait pu être l'inverse). Tous les caractères en x sont donc plésiomorphes et les caractères en x' sont apomorphes.

La comparaison extra-groupe

La principale manière de polariser un caractère est l'utilisation d'un groupe externe ou extra-groupe (*outgroup*). On estime que tout état de caractère présent en dehors du groupe interne (ou *in-group*) est antérieur au groupe interne car déjà présent ailleurs. On va donc considérer ce caractère comme primitif. Tout état de caractère présent en dehors du groupe d'étude est donc considéré comme primitif et tout état qui n'est présent que dans le groupe d'étude est considéré comme dérivé, car propre au groupe d'étude.

Cette règle s'explique par le fait que l'on admet qu'un caractère partagé par divers taxons est hérité de l'espèce ancestrale de ces taxons.

Au lieu d'utiliser cette procédure pour chaque caractère, on utilise un groupe entier en considérant tous ses caractères. Ce groupe doit être considéré a priori comme externe du groupe d'étude. Le groupe externe ne doit pas être trop éloigné du groupe d'étude sinon les caractères ne pourront même pas être comparés. A l'inverse, s'il est trop proche le groupe externe peut en fait faire partie du groupe interne. De plus, si certains états de caractères du groupe externe sont homoplasiques avec ceux du groupe interne, la polarisation de certains caractères peut être fautive. Le bon choix du groupe externe est crucial et un mauvais choix peut entraîner des incohérences. Le groupe externe permet alors d'enraciner l'arbre en l'orientant.

L'ordination chronologique

Des données d'ordre historique peuvent permettre d'ordonner certains caractères. Le caractère qui existe dans un contexte historique ancien est primitif, tandis que celui qui existe dans un contexte récent est dérivé. On se souviendra pourtant que l'ancienneté historique et l'état primitif des caractères ne sont véritablement liés que dans le cas de relations directes d'ancêtre à descendant.

La congruence

Le terme est adopté dans son sens mathématique d'identité géométrique. Si deux taxons sont proches parents, et donc considéré comme deux groupes frères, on peut s'attendre à reconnaître d'autres traits que l'on n'observe pas ailleurs. Le principe de congruence renvoie au principe de parcimonie.

Le clade comme unité de classification

Julian Huxley, l'homme de la « synthèse moderne » formalisa en 1957 les différents types de processus évolutifs en distinguant anagenèse fondée sur l'analogie et cladogénèse fondée sur l'homologie et leurs produits taxonomiques : grades et clades.

L'unité anagénétique « grade » tient des processus, en l'occurrence une conception du progrès évolutif ; l'unité monophylétique « clade » tient de la structure (la généalogie), une façon de faire réapparaître le dualisme entre théorie des processus et théorie de la structure.

Homologie versus homoplasie

Dans la perspective évolutionniste, les critères d'identification de l'homologie restent des critères structuraux qui tiennent de l'observation empirique. Des organes observés dans divers taxons sont considérés comme homologues dans la mesure où ils entretiennent entre eux des connexions identiques. Identifié par Etienne Geoffroy Saint-Hilaire, ce principe sera repris par Richard Owen. Deux organes sont homologues si, quelles que soient leurs formes ou fonctions, ils ont les mêmes connexions avec d'autres organes.

Cette homologie est ensuite interprétée comme une hypothèse phylogénétique *sensu lato*: ce qui est similaire est dû à la descendance. Les caractères homologues partagés par différentes espèces sont hérités d'une espèce ancestrale commune : ils ont une histoire propre.

À l'inverse, l'homoplasie, terme introduit par Lankester en 1870, caractérise, au sein d'un ou de plusieurs taxons, des ressemblances entre caractères n'étant pas hérités d'un ancêtre commun. Dans ce cas le même caractère peut apparaître de façon indépendante dans deux lignées.

Différencier l'homoplasie de l'homologie est une tâche compliquée. En effet rien ne nous assure que le caractère supposé homologue l'est effectivement et que ce n'est pas de l'homoplasie.

L'hypothèse de base est que le même caractère observé chez deux taxons indique une relation de parenté. Dans une perspective phylogénétique on cherche donc à maximiser l'homologie (ou à minimiser l'homoplasie, ce qui constitue deux démarches équivalentes). Les caractères n'apparaissant qu'une fois dans le cladogramme résultant de l'analyse de parcimonie seront considérés comme effectivement homologues et donc hérités d'un ancêtre commun. L'homologie n'est donc plus seulement une homologie déduite du principe des connexions mais une homologie de parenté ou de descendance.

À l'inverse, certains caractères peuvent se révéler homoplastiques une fois l'arbre reconstitué. Ces caractères impliquent au moins deux pas évolutifs ou deux hypothèses de transformation dans le cladogramme résultant. L'homoplasie est donc une hypothèse d'homologie primaire rejetée.

Il existe en fait deux types d'homoplasie: la convergence évolutive et la réversion, décelables entre elles seulement sur l'arbre enraciné le plus parcimonieux.

1. La convergence indique que le même caractère (au sens des connexions et non au sens évolutif) est apparu au moins deux fois dans le groupe considéré. Cette dernière s'effectue sur plusieurs axes distincts.
2. À l'inverse, la réversion est la perte secondaire d'un caractère, c'est-à-dire le retour à l'état primitif. Une réversion est une homoplasie qui s'effectue sur un même axe présentant une suite de transformations. L'homoplasie n'est donc pas un caractère hérité par un ancêtre commun et ne nous renseigne pas sur les relations de parenté. Un groupe formé sur la base d'homoplasies est appelé groupe polyphylétique et constitue un grade.

Clades

Un clade est un groupe de taxons formant un groupe monophylétique, une totalité de descendance, un ancêtre commun et tous ses descendants. Dans un clade, qui représente une unité évolutive, les taxons sont plus proches entre eux qu'avec de tout autre taxon. Il est à noter qu'un clade représente uniquement une unité évolutive.

Grades

Un grade est un rapprochement de taxons reposant sur d'autres critères (ressemblance générale, somme de modifications adaptatives, etc.). Un grade illustre donc un stade de progrès, de perfection évolutive. On dira aujourd'hui avec moins d'emphase, qu'un grade est un ensemble d'êtres vivants qui partagent un même stade ou niveau évolutif général.

Un grade n'est pas nécessairement un clade et peut-être acquis indépendamment par divers groupes et être polyphylétique ou paraphylétique : ses membres sont les descendants de diverses formes ancestrales et ne sont pas unis par des liens phylogénétiques exclusifs. Ils ne sont pas tous les descendants d'une espèce ancestrale et, par conséquent, n'ont pas d'histoire propre. Il n'est pas un groupe naturel.

Un grade peut être paraphylétique. Les rapprochements sont effectués sur la seule base de caractères primitifs donc plésiomorphes. Ce n'est donc pas une totalité de descendance puisque les organismes portant les états dérivés en sont exclus.

Un grade peut être polyphylétique. Les rapprochements sont effectués sur la base de caractères analogues donc d'homoplasies.

Il est donc important en cladistique de bien séparer caractère plésiomorphe et apomorphe pour déterminer si un groupe est monophylétique ou non. Dans un groupe paraphylétique, certains taxons sont plus proches d'autres taxons hors du groupe que de taxons au sein du groupe paraphylétique.

Taxons

Le taxon est l'unité d'analyse phylogénétique regroupant un certain nombre de traits jugés pertinents. En cladistique classique le taxon est l'espèce biologique ou de son équivalent paléontologique supposé. Pour notre présent propos, le taxon peut être un type postulé de société.

La construction phylogénétique ne passe pas par une recherche des relations d'ancêtre à descendant. Deux taxons sont toujours apparentés. Il est plus intéressant, afin de construire une hypothèse scientifique, de chercher à savoir si deux taxons parmi trois sont proches parents. Il est toujours possible d'émettre ce type d'hypothèse et il n'y a, à ce stade de l'analyse, aucune recherche de relations directes d'ancêtre et de descendants.

Dans cette perspective il est possible par exemple d'émettre l'hypothèse que les sociétés C et D sont étroitement apparentées si l'on met en évidence le partage d'un caractère hérité d'une société ancestrale propre et qui, en conséquence, n'est pas présent dans la société B. A aucun moment n'est recherché un lien direct d'ancêtre à descendant parmi les trois sociétés B, C et D.

Le partage de caractères dérivés, éventuellement insignifiants, est seul indicateur d'une proche parenté et non celui de caractères primitifs, même nombreux, même spectaculaire. Similitude globale et similitude due à la filiation ne sont pas nécessairement superposables.

Deux sociétés sœurs, c'est-à-dire deux taxons Y et Z dont on pense qu'ils sont étroitement apparentés, descendent d'une société ancestrale exclusive, présentent rarement le même degré évolutif global vis-à-vis de leur ancêtre. L'une des deux est souvent plus éloignée morphologiquement de l'ancêtre en raison d'un plus grand nombre de traits dérivés propres : c'est la notion de divergence des évolutionnistes. Il n'en reste pas moins vrai que du point de vue de l'analyse des caractères, la société Y (au cas où elle n'aurait pas évolué) n'est pas recherchée comme un ancêtre potentiel, même si, a priori, il apparaît que ses caractères ne sont pas différents de ceux de l'ancêtre commun.

Le cladogramme ne nous montre donc pas des sociétés évoluées émergeant de sociétés primitives mais une succession de groupes reliés deux à deux, les groupes frères. Une évolution sans ancêtre ? Pas tout à fait : Les nœuds du cladogramme représentent les ancêtres et leurs descendants au cours de temps : mais ces ancêtres sont hypothétiques, reconstruits et non directement observés. Le paradoxe tient aussi à la définition de l'ancêtre. Selon le cladisme l'ancêtre est un ancêtre au sens strict : la plus petite unité évolutive identifiable : société, populations, groupes d'individus.

Principe de parcimonie

La philosophie derrière les méthodes de parcimonie se fonde sur un principe attribué à Guillaume d'Ockham, selon lequel les entités ne doivent pas être multipliées par delà ce qui est nécessaire (« *Entia non sunt multiplicanda praeter necessitatem* »). Pour faire un choix entre deux arbres, on va donc considérer que le plus court, i.e. l'arbre avec le moins de pas évolutifs, est l'arbre représentant l'hypothèse phylogénétique la plus plausible.

Le principe de parcimonie est plus connu en France sous le nom de principe d'économie d'hypothèse. La meilleure explication d'un fait ou d'un ensemble de faits est celle qui nécessite le minimum d'hypothèses, ou plus trivialement : ne pas faire compliqué quand on peut faire simple. Il convient donc de répondre à la question en quoi la construction des arbres phylogénétiques s'accompagne de ce principe. La construction phylogénétique consiste en effet à édifier un arbre qui contient le plus petit nombre de pas évolutifs : c'est l'arbre le plus « court », le plus parcimonieux.

Il est important de bien saisir à ce propos que le fondement du système phylogénétique est l'analyse des caractères pris chacun individuellement : les caractères sont libres d'évoluer indépendamment les uns des autres.

On distingue dans cette perspective trois modèles de parcimonie, selon la considération des convergences et des réversions :

- La parcimonie de Wagner, où convergences et réversions sont acceptées
- La parcimonie de Camin-Sokal admet les convergences mais pas les réversions.
- La parcimonie de Dollo accepte les réversions mais pas les convergences.

La descendance avec modification est un modèle parcimonieux. Le principe de parcimonie induit les critères mêmes d'identification du sens des transformations des caractères, autrement dit la polarité.

L'observation d'un même état dérivé dans deux sociétés ne doit pas être considéré a priori comme due à une convergence. Le principe de parcimonie, ou de congruence, prend précisément toute son importance lorsque la distribution de plusieurs caractères soulève des contradictions, autrement dit lorsqu'il est nécessaire d'admettre une part d'homoplasie.

Pour expliquer la présence d'un même caractère dérivé chez deux taxons, la synapomorphie représente l'hypothèse la plus simple (un pas). Les hypothèses auxiliaires, ad hoc,

correspondent aux homoplasies. La théorie des relations de parenté qui renferme le moins d'hypothèses ad hoc (illustrées par le cladogramme) est préférée à celle qui est plus chargée d'hypothèses. Le meilleur arbre est celui où chacun des traits n'apparaît qu'une fois et, à défaut (car il y a toujours des homoplasies), c'est celui où chacun des traits apparaît un nombre de fois minimal.

C'est parce qu'il est fondé sur le principe de parcimonie que le cladogramme peut être testé et soumis à réfutation. Par son ouverture au test, sa réfutabilité, le cladogramme est une construction « poppérienne », pourrait-on dire, dont la condition logique de la mise en œuvre est le principe de parcimonie.

On a pu soutenir que si un seul caractère était en contradiction avec les autres, le cladogramme était réfuté. En fait, ce n'est pas le cladogramme qui est réfuté, mais le caractère en question, caractère pouvant être soumis à discussion, à élimination ou à un nouveau codage.

Structures et mécanismes : l'explication causale

Les pratiques cladistiques ont dissocié définitivement structure (le *pattern* des anglophones) et processus (les mécanismes de notre terminologie). Point n'est besoin de connaître les processus, les mécanismes, pour ordonner les produits de l'évolution en un système hiérarchique, c'est-à-dire un arbre.

Il est donc important de bien distinguer les deux concepts différents, le *pattern* et le *process*. La cladistique permet de comprendre le *pattern* ou la structure des sociétés actuelles ou passées, c'est-à-dire leurs relations de parentés, tandis que l'évolution répond à la notion de processus évolutif ou *process*, c'est-à-dire comment les caractères des sociétés évoluent et sont acquis (par évolution, révolution, invasion, influence, etc.). Dès lors les cladogrammes représentent uniquement des classifications entre sociétés ou taxons. Il est cependant évident que les caractères, qui permettent de faire des clades, sont le résultat de mécanismes évolutifs mais la cladistique ne cherche, provisoirement, ni à les expliquer, ni à les connaître. Seuls les liens de parenté sont recherchés en cladistique. Cependant le cladogramme résultant de l'analyse permet d'interpréter l'évolution des clades et des caractères. En effet, le clade étant une unité évolutive, connaître la structure de parenté des taxons permet d'interpréter la distribution des caractères au cours du temps, et donc l'évolution des caractères impliqués.

Structures et scénarios : l'histoire retrouvée

Le cladogramme peut être facilement transformé en un arbre phylogénétique traditionnel : il suffit d'y inscrire la dimension historique des sociétés étudiées, passées et présentes, mêlées et analysées selon les mêmes critères. Il peut même être transformé en scénario évolutif.

La société souche, assimilable dans une certaine mesure au groupe ancestral évolutionniste, est un pis-aller. Nécessairement paraphylétique, c'est un groupe retenu provisoirement, destiné à être démembré à l'occasion de découvertes ultérieures, et non le nec plus ultra de l'analyse phylogénétique.

L'arbre phylogénétique cladistique ne contient qu'un seul type d'information : la séquence de branchement, qui n'est autre que l'histoire de l'apparition des lignées. La dimension chronologique est implicite : l'arbre illustre l'ancienneté relative des ancêtres respectifs des groupes subordonnés. Mais l'échelle de temps peut ne rester qu'implicite : La structure des relations de parenté est exprimée, mais non l'extension dans le temps historique.

Une bonne théorie de l'évolution naîtra néanmoins d'une bonne systématique, et non l'inverse. Rien n'empêche de transformer le cladogramme en scénario évolutif ou phylogramme. Seuls les groupes ancestraux manquent à l'appel d'un arbre entièrement résolu.

A ce titre le cladisme s'inscrit au départ dans l'approche évolutionniste de la systématique et reprend à son compte la prédiction de Darwin : nos classifications deviendront des généalogies. La logique du cladogramme est indissociable de son interprétation évolutive : les groupes monophylétiques sont les seuls à posséder une origine et une histoire propres. Les artefacts taxinomiques, groupes polyphylétiques, paraphylétiques et autres groupes ancestraux, n'ont aucune histoire propre et ne répondent à aucune logique de distribution de caractères.

Sociétés ancestrales

En cladistique une société ne peut pas être primitive: il n'y a pas de groupe ancestral. Si un clade contient plusieurs groupes, c'est que chacun de ces groupes a des caractères dérivés qui les distinguent des autres, ce qui implique qu'un groupe ancestral ne peut pas exister en cladistique. Une société ancestrale est définie par des caractères primitifs donc plésiomorphes par rapport à ses descendants, elle est donc forcément paraphylétique. En cladistique l'ancêtre est une hypothèse et n'est pas identifié, ce qui, évidemment, n'implique pas qu'il n'existe pas, mais qu'il ne peut pas être défini en tant que taxon ou société mais simplement en tant qu'hypothèse d'un ensemble de caractères primitifs (on parle alors de morphotype ancestral).

Le modèle hennigien dénie toute pertinence phylogénétique aux caractères primitifs. Il implique que tout groupe construit a contrario à partir de traits primitifs n'a pas d'existence phylogénétique. A partir de telles prémisses, force est de conclure que les groupes ancestraux paraphylétiques n'existent pas, à l'inverse de ce que voudrait montrer les arbres néodarwiniens. Ils ne partagent que des caractères primitifs connus ailleurs, mais pas de caractère dérivé propre inconnus ailleurs. En conséquence de tels groupes ne sont pas des communautés de descendance et n'ont pas d'histoire propre : les événements évolutifs, jalonnant et guidant leur histoire, ne leurs sont pas exclusifs. Dépourvu de groupe ancestral, l'arbre phylogénétique cladistique a une forme comparable à celle d'un phénogramme produit par les techniques phénétiques, sauf qu'il possède un point de départ, une racine, symbolisant l'ancêtre commun le plus lointain du groupe. Cette racine est « le morphotype ancestral hypothétique ».

Il se peut néanmoins que des sociétés historiques répondent aux attributs de ce morphotype et soient connus à une époque compatible avec l'hypothèse de relation d'ancêtre à descendants. On admettra dans ce cas avoir découvert un véritable ancêtre.

Références

BALANDIER G. 1974. *Anthropo-logiques*. Paris : PUF.

DARLU P., TASSY P. 1993 (3ème édition). *La reconstruction phylogénétique du vivant*. Version électronique.

GALLAY A. 1986. *L'archéologie demain*. Paris : Belfond (Belfond sciences)

GALLAY A. 2006 (rééd. 2011). *Les sociétés mégalithiques : pouvoir des hommes, mémoire des morts*. Lausanne : Presses polytechniques et univ. romandes. (Le savoir suisse : histoire ; 37).

GALLAY A. 2011a. *De mil, d'or et d'esclaves : le Sahel précolonial*. Presses polytechniques et univ. romandes. (Le Savoir suisse : histoire ; 72).

- GALLAY A, 2011b. *Pour une ethnoarchéologie théorique*. Paris : Actes Sud/Errance.
- TESTART A. 2005. *Éléments de classification des sociétés*. Paris : Errance.
- GOULD S. J. 2002. *La structure de la théorie de l'évolution*. Paris : Gallimard. NRF essais.
- HENNIG W. 1950. *Grundzüge einer Theorie der phylogenetische Systematik*. Berlin : Deutscher Zentralverlag (Deutsche entomologische Institut Berlin-Friedrichshage).
- HENNIG W. 1966. *Phylogenetic systematics*. Chicago : University of Illinois press.
- HUXLEY J. 1957. The three types of evolutionary process. *Nature*, 180, p. 454-457.
- MAYR E. 1963. *Animal species and evolution*. Cambridge (MA) : Harvard university y press
- TASSY P. 1991. *L'arbre à remonter le temps : les rencontres de la systématique et de l'évolution*.
- TESTART A. 2005a. *Éléments de classification des sociétés*. Paris : Errance.