

Chapitre 18

Classer la biodiversité

Cours

PLAN DU CHAPITRE

- 1 Principes d'établissement d'une classification
- 2 L'approche phylogénétique cladistique

ZOOM

- 1 Caractères utilisés pour classer les organismes
- 2 Phénogramme ou cladogramme

INTRODUCTION

La **biodiversité** recouvre la multiplicité des organismes d'une espèce, d'un écosystème et des différents écosystèmes de la biosphère. L'étude des fossiles montre quant à elle des changements de la biodiversité au cours des temps géologiques. Chercher à classer la diversité des êtres vivants actuelle et passée ne revient pas seulement à les trier ou à les ranger ; cela suppose d'établir entre eux des relations de parenté (relations phylogénétiques).

➔ **Quels sont les principes d'établissement d'une classification ?**

➔ **En quoi consiste l'approche phylogénétique cladistique actuellement reconnue ?**

1 Principes d'établissement d'une classification

Diverses modalités de classement existent en biologie, selon les objectifs poursuivis. Elles reposent sur différentes approches de la notion de ressemblance.

1.1 Classification ou tri ?

a) La taxonomie, discipline visant à décrire et à identifier la biodiversité

La **taxonomie** (ou taxinomie) a pour objet la description des êtres vivants permettant d'établir leur appartenance à des espèces déjà répertoriées, ou d'en définir de nouvelles. Elle a aussi pour objet de pouvoir ranger chacune de ces espèces parmi des ensembles (ou **taxons**) organisés de manière emboîtée (figure 18.1).

Karl von Linné a défini la première classification hiérarchique. Chaque espèce y est nommée de manière binominale, par un nom de genre suivi du nom d'espèce, ces deux termes étant empruntés au latin. Par la suite, divers rangs taxonomiques d'ordres supérieurs à celui de l'espèce ont été définis ; ils sont désignés par des termes latins débutant par une majuscule et auxquels sont attachés des noms usuels en français (figure 18.1).

La distinction de taxons de rangs différents repose sur le principe de « **subordination des caractères** ». Par exemple, puisque la présence d'une chorde est un caractère partagé par un plus grand nombre d'organismes que la présence de poils, le taxon des chordés est de rang supérieur à celui des mammifères (*i.e.* il l'inclut).

DÉCOUVERTE 1

Taxonomie et classification : un bref historique

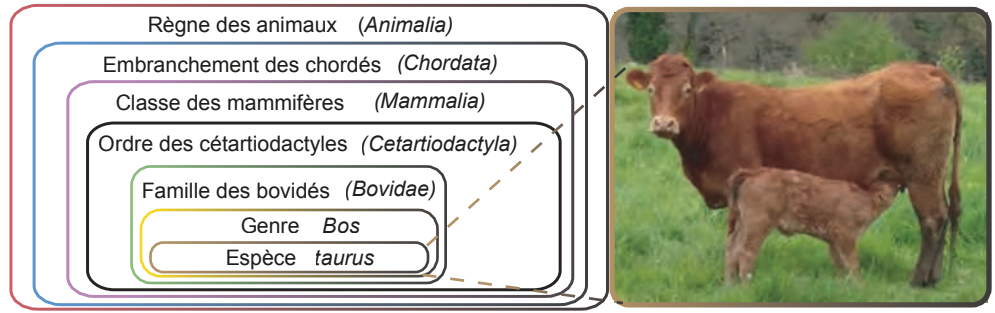


Figure 18.1 Position taxonomique de la vache et du veau.

Sont ici représentés les sept rangs formels définis par Linné, avec leur nom usuel en français et la terminologie internationale en latin entre parenthèses. D'autres pourraient venir s'intercaler comme le sous-embranchement des vertébrés (*Vertebrata*) entre chordés et mammifères. L'espèce *Bos taurus* désigne aussi bien la vache que le veau ou le taureau.

b) Différents critères de tri, en fonction des objectifs poursuivis

• **Identification des espèces par une clé de détermination**

La taxonomie traditionnelle repose sur une approche simple, encore utilisée aujourd'hui pour identifier les espèces et dresser des inventaires de la biodiversité. Elle ne traduit aucun lien de parenté. Les **clés de détermination** reposent en général sur un mode de questionnement et de progression dichotomique (présence ou absence d'un caractère ou d'un lot de caractères) qui reflète la hiérarchie de la nomenclature linnéenne. C'est par exemple le cas des flores dont l'utilisation a déjà été abordée.

• **Identification des espèces par le **barcodage**.**

ZOOM 1
Caractères utilisés pour classer les organismes
Voir TP 11, § 2

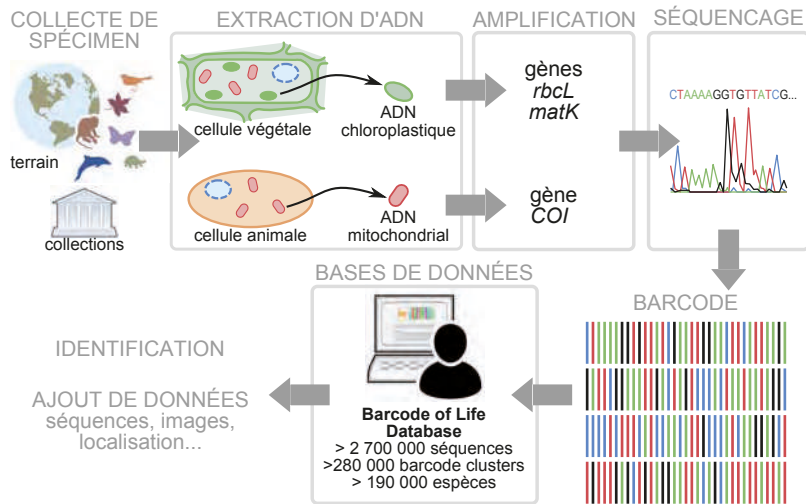


Figure 18.2 La procédure du barcodage.

Le projet iBOL (*international Barcode of Life*) débuté en 2010 vise à l'établissement d'une base de données d'un demi-million d'espèces « barcodées ». Chez les animaux, la séquence de référence est celle du gène codant la sous-unité I de la cytochrome-oxydase mitochondriale (gène *COI*). Chez les plantes dont le génome mitochondrial est moins variable que celui des animaux, deux séquences plastidiales sont utilisées conjointement. Chez les bactéries, ce sont les séquences d'ADN ribosomique 16S qui sont utilisées.

Voir chapitre 17,
figure 17.4

Voir chapitre 17,
figure 17.1

Le barcodage consiste à caractériser une espèce par une séquence d'ADN donnée, le barcode (par analogie avec les codes-barres présents sur des objets de la vie courante). La séquence de référence retenue est d'origine mitochondriale chez les animaux et chloroplastique chez les végétaux (figure 18.2). Ce choix est justifié par la grande variabilité interspécifique de ces séquences, par la présence d'un grand nombre de leurs copies au sein des cellules et par un isolement plus facile que pour une séquence nucléaire. Cette méthode trouve aussi de très nombreuses applications en écologie.

• Classification en groupes fonctionnels

D'autres « classifications » peuvent être établies pour répondre à des objectifs variés, comme par exemple celle de **groupes fonctionnels** dans un écosystème.

En résumé, le terme « classification » recouvre parfois une simple démarche de tri des êtres vivants, selon divers critères ; de telles méthodes ne cherchent pas d'interprétation évolutive des regroupements effectués ou des différences constatées.

c) La systématique, discipline visant à classer des taxons

Le travail du systématicien consiste à substituer à un tri de la biodiversité en groupes emboîtés et fondés sur le partage de caractères communs, une **classification phylogénétique** visant à reconstituer les liens de parenté entre organismes vivants et éventuellement fossiles. Ces liens de parenté peuvent être représentés par des **arbres phylogénétiques**.

Toutes les classifications s'appuient sur les **mêmes caractères** (morphologiques, anatomiques, fonctionnels, ou moléculaires) et utilisent des ressemblances entre les êtres vivants. Leurs différences viennent de l'interprétation qui est faite de ces ressemblances et notamment du choix de celles qui sont retenues comme critère d'apparement.

Différentes approches se sont succédé au cours du temps, jusqu'à la systématique phylogénétique actuelle qui s'est développée au cours de la seconde moitié du XX^e siècle dans la suite des idées évolutives darwiniennes.

1.2 Différentes approches pour établir une classification

On distingue ainsi trois écoles différentes de systématique qui divergent principalement dans leur façon d'établir un apparement.

a) La systématique phénétique

Cette école établit des « classifications » basées sur une **ressemblance globale entre taxons** (phénétique renvoie au terme grec « *phainein* » = paraître). L'approche phénétique s'appuie sur des matrices établies qui consignent les degrés de similitude entre caractères, pour chaque couple de taxons. À partir de ces données, on peut construire des figures arborescentes nommées **phénogrammes** qui représentent le **degré de similarité (DGS) entre des êtres vivants**, sans que celui-ci ne soit forcément corrélé à leur parenté. Ainsi, les ramifications d'un phénogramme rendent compte des différences (ou divergences phénétiques) entre les taxons étudiés, mais n'établissent pas obligatoirement des relations de parenté, car toutes les ressemblances ne sont pas le signe de liens phylogénétiques.

Après un vif succès dans les années 1970, surtout en raison de son formalisme mathématique, la systématique phénétique disparaît peu à peu ; il ne subsiste aujourd'hui qu'une **méthode** phénétique utilisée en phylogénie moléculaire (méthode du *Neighbor-Joining*, « NJ », voir [tableau 18.3](#)), qu'il convient de ne pas confondre avec une systématique phénétique.

b) La systématique biologique ou fonctionnelle

Cette école de pensée, nommée **école gradiste**, repose sur l'idée d'une évolution qui serait complexifiante et progressiste. Des ressemblances morpho-anatomiques et embryologiques, sont censées rendre compte de liens phylogénétiques ; elles sont réunies dans une première

ZOOM 2

Phénogramme ou
cladogramme

idée de l'**homologie** (homologie primaire). On identifie un **grade** comme un groupe d'êtres vivants apparentés parce que partageant un même plan général d'organisation. Les différences sont alors interprétées comme des « sauts » d'une évolution supposée complexifiante. Une des principales critiques adressées à cette approche repose sur le choix arbitraire des caractères sur lesquels fonder la notion de saut évolutif (figure 18.3).

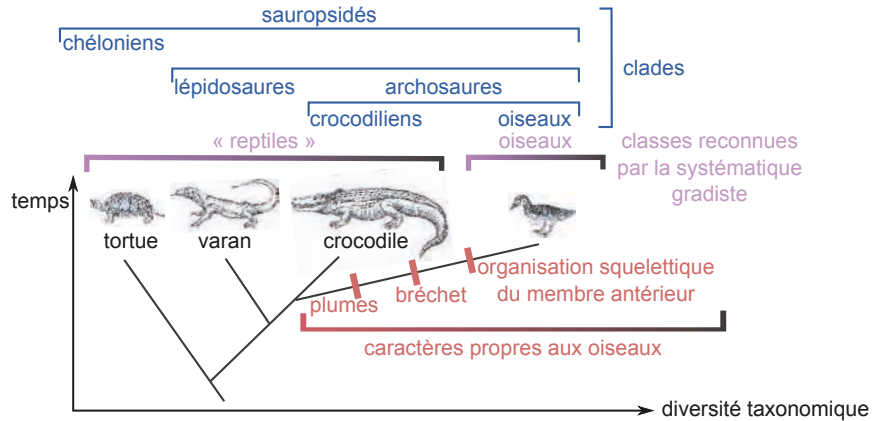


Figure 18.3 Classification gradiste et cladiste des sauropsidés.

La systématique gradiste reconnaît le statut de classe pour les reptiles et pour les oiseaux, en considérant ces derniers comme le fruit d'un saut évolutif justifiant leur séparation par rapport aux reptiles (nombreuses adaptations liées au vol chez les oiseaux : plumes, bréchet, ailes). La systématique cladiste privilégie quant à elle des ressemblances morpho-anatomiques entre oiseaux et crocodiliens, pour les rassembler au sein du clade des archosaures et les distinguer de celui des lépidosaures. Si elle reconnaît également les oiseaux comme un clade au sein des archosaures, elle ne permet pas de faire de même pour les reptiles. En cladistique, seul le groupe monophylétique a une valeur (voir § 1.2c).

c) La systématique phylogénétique ou cladistique

La cladistique se développe grâce aux travaux de W. Hennig à partir des années 1950. Elle désigne à la fois une école de pensée et une méthode de construction phylogénétique (voir § 2). Ce mode de classification ne retient que des **ressemblances particulières**, ou **apomorphies** pour classer des organismes dans un **groupe monophylétique**. Ce sont des caractères nouveaux partagés (d'où le terme de **synapomorphies** utilisé pour les nommer) par plusieurs organismes représentatifs de différents taxons, et considérés comme hérités d'un ancêtre commun. Ils sont aussi nommés homologies vraies (ou **homologies secondaires**, figure 18.4a), et constituent les indices d'apparement qui permettent d'établir des groupes monophylétiques, ou clades. Un groupe **monophylétique** comprend un ancêtre commun et l'**ensemble** de ses descendants, partageant au moins une apomorphie. C'est le cas du clade des oiseaux qui partagent des apomorphies telles que les plumes, le bréchet (os de la cage thoracique sur lequel s'insère des muscles du vol), et des pièces osseuses particulières dans leur membre antérieur (figure 18.3). L'ancêtre commun est défini par la collection des synapomorphies.

Remarque

Certains membres d'un clade ont pu secondairement perdre, ou subir une transformation d'une ou plusieurs des apomorphies caractéristiques sans que cela ne remette en cause leur appartenance à ce clade et donc l'existence de celui-ci. Par exemple, les plastes sont absents chez les oomycètes (agents de maladie des plantes, comme le mildiou) sans que cela ne remette en cause leur appartenance au clade des straménopiles dont une synapomorphie est la présence de plastes.

Voir ouvrage de
2^e année

- Un groupe **paraphylétique** est fondé sur le partage de caractères non innovants, qualifiés de plésiomorphies (ou **symplesiomorphies** pour signifier qu'un caractère est partagé par un groupe d'organismes). Un groupe paraphylétique, comme celui des reptiles au sein des sauropsidés, est un regroupement artificiel dans lequel il manque certains descendants de l'ancêtre commun, les oiseaux (figures 18.3 et 18.4b).
- Quant au groupe **polyphylétique**, il est établi sur des ressemblances entre des caractères apparus indépendamment dans des branches évolutives distinctes, plus ou moins éloignées. Dans un groupe polyphylétique, il manque l'ancêtre commun lui-même. Ceci n'apparaît pas forcément sur la figure 18.4 qui est très simplifiée. L'ancêtre commun d'un groupe polyphylétique est aussi celui d'un très grand nombre de taxons (« poly » : nombreux).

Une ressemblance entre individus fondée sur le partage de deux états d'un caractère apparu indépendamment est nommée **homoplasie** (figure 18.4c et d). Une homoplasie peut avoir pour origine des acquisitions indépendantes d'états dérivés, se ressemblant structurellement et fonctionnellement. Il en est ainsi de la présence d'ailes chez les oiseaux et chez des mammifères comme la chauve-souris, qui sont le fruit d'une **convergence évolutive**. Compte tenu de leur proximité évolutive, on parle d'homoplasie de convergence. Par contre l'aile des insectes et celle des oiseaux caractérise bien une ressemblance fonctionnelle (l'aptitude au vol) mais entre des organismes éloignés sur le plan phylogénétique ; cette ressemblance est donc interprétée comme une « convergence fonctionnelle » et on lui réserve alors le terme d'**analogie**. Une homoplasie peut aussi résulter d'une ou plusieurs réversions faisant réapparaître des états ancestraux (nageoires paires du maquereau et du dauphin, par exemple).

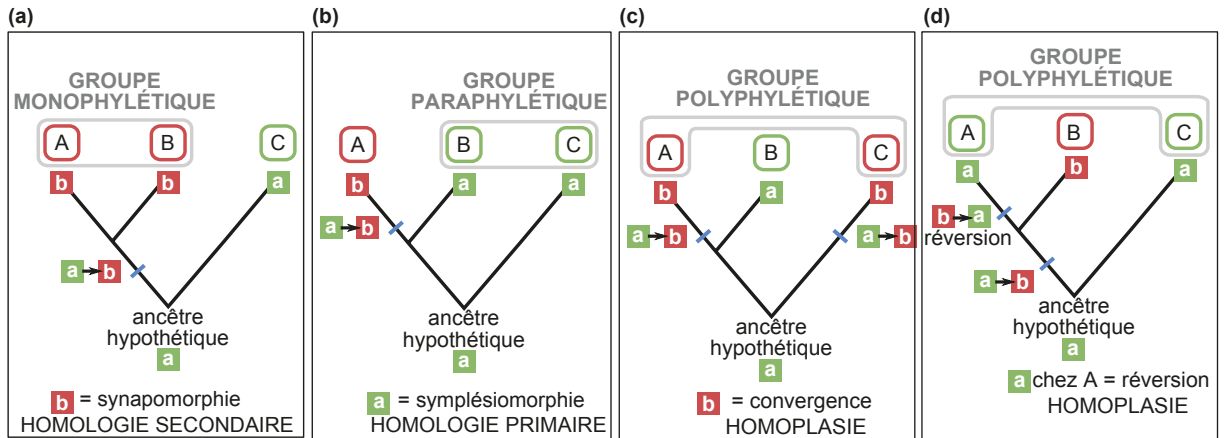


Figure 18.4 Différents types de ressemblances : homologie secondaire (cas a), homologie primaire (cas b, état de symplesiomorphie) et homoplasies (cas c et d).

(Modifiée d'après Lecointre et Le Guyader, 2009).

2 L'approche phylogénétique cladistique

La **phylogénie** est la discipline scientifique dont l'objectif est de reconstruire l'histoire des organismes vivants. La démarche phylogénétique s'appuie sur des hypothèses représentées par des arbres que l'on teste. Un arbre phylogénétique donné représente une **hypothèse** retraçant l'histoire d'un groupe particulier d'organismes et traduit leurs relations de parenté supposées : « qui est plus proche de qui ? » ou « quels organismes ont une ascendance commune ? ». Contrairement à un arbre généalogique, un arbre phylogénétique ne permet pas de répondre à la question : « qui descend de qui ? ». La phylogénie ne peut proposer que **des scénarios hypothétiques** sur l'évolution des organismes ; ces scénarios dépendent des données disponibles.

Voir chapitre 1,
Zoom 1

ZOOM 1

Caractères utilisés
pour classer les
organismes

Voir TP1, figure TP1.2

DÉCOUVERTE 1

Taxonomie
et classification :
un bref historique

2.1 Construire un arbre par raisonnement cladistique

Le raisonnement cladistique est une façon d'établir une phylogénie. Il conduit à une représentation codée de son résultat, appelée cladogramme.

a) Recenser les caractères permettant de poser une hypothèse d'homologie

Contrairement à ce que postule l'approche phénétypique, apparemment et ressemblance entre organismes ne sont pas forcément corrélés. Tous les caractères ne sont pas utilisables pour établir une phylogénie : certains (couleurs des yeux) sont liés à la diversité phénotypique intraspécifique, d'autres ne sont tout simplement pas comparables d'un groupe à un autre (ailes des insectes et des oiseaux).

Les caractères utilisables sont ceux pour lesquels il est possible non seulement de constater une ressemblance structurale et fonctionnelle mais aussi de supposer qu'elle témoigne d'une parenté. On parle alors d'une **homologie primaire**, hypothèse de parenté qu'il est possible de poser à l'aide des données suivantes.

- Le principe des **connexions anatomiques** : une hypothèse d'homologie peut être posée pour deux caractères situés dans une même position relative chez deux organismes, et ce, quelles que soient leurs formes et leurs fonctions. C'est par exemple le cas des divers segments du membre chiridien des vertébrés tétrapodes (stylopode, zeugopode, autopode) que l'on peut retrouver dans des structures aussi différentes que le bras de l'homme, la patte antérieure de la vache ou l'aile d'un oiseau.
- Une origine **embryologique** commune : par exemple, les trois os de l'oreille moyenne (marteau, étrier, enclume) des mammifères dérivent embryologiquement des mêmes pièces squelettiques que celles affectées à l'articulation de la mâchoire des oiseaux. Une homologie primaire peut donc être proposée pour les os de l'oreille moyenne des mammifères et les os mandibulaires des oiseaux.
- Des données **paléontologiques** : ce changement d'affectation d'os entre la mandibule des oiseaux et l'oreille moyenne des mammifères est illustré chez divers vertébrés fossiles du Permian-Trias qui paraissent comme des étapes intermédiaires de ce transfert. C'est un argument de plus pour postuler l'homologie.
- Des données **moléculaires** : on considère que deux séquences protéiques ou nucléotidiques sont homologues si plus de 20 % des différents sites de leurs séquences alignées sont occupés par le même monomère, acide aminé ou nucléotide.

Il conviendra, dans la suite de la démarche, de tester ces hypothèses de parenté.

b) Recenser les caractères pouvant être polarisés

Pour être utilisable dans la démarche cladistique, un caractère doit aussi posséder les propriétés suivantes.

- **Exister sous au moins deux états** : par exemple, « présence ou absence » de colonne vertébrale chez les métazoaires ; « présence ou absence » de tissus conducteurs chez les embryophytes.
- **Être polarisable** : parmi les différents états d'un caractère, on doit pouvoir distinguer :
 - un état ancestral (ou **plésiomorphie**) ;
 - un (ou plusieurs) état(s) dérivé(s) (ou **apomorphie**) : cet état nouveau est acquis par transformation de l'état ancestral, chez un ancêtre d'un ou plusieurs membres du groupe étudié.

Plésiomorphie et apomorphie ne constituent que des notions toutes relatives (figure 18.6).

Définir l'état ancestral et l'état dérivé des caractères étudiés suppose de disposer de critères de polarisation. Deux méthodes de polarisation sont généralement utilisées.

- Le recours au **développement embryonnaire** utilise l'idée de Haeckel suivant laquelle l'ontogénèse récapitule la phylogénèse : si un état du caractère précède l'autre lors du développement d'un organisme, alors il est considéré comme ancestral. Cependant cette

récapitulation est loin d'être universelle ; elle est même considérée aujourd'hui comme un phénomène peu fréquent.

- On peut aussi se référer à un **groupe externe** (encore appelé « extra-groupe »), choisi sur la base de travaux antérieurs, à l'extérieur du groupe de taxons dont on souhaite étudier les parentés. Ce groupe externe présente tous les états ancestraux des caractères choisis pour mener l'étude cladistique. On dit qu'il permet d'**enraciner** l'arbre qui classera les uns par rapport aux autres les taxons étudiés.
- **Être suffisamment représenté sous ses deux états.** Un caractère dérivé présent chez un seul membre du groupe étudié n'apporte rien sur son apparentement précis vis-à-vis des autres. De même un caractère présent à l'état dérivé chez tous les membres du groupe permet juste de les classer par rapport au groupe externe.

L'établissement d'une phylogénie à partir d'un groupe de taxons sera d'autant plus abouti qu'un nombre important de caractères homologues (au sens de l'homologie primaire, *i.e.* homologie supposée) aura été pris en compte.

c) Organiser l'information pour faciliter le raisonnement

• Construire une matrice de caractères

Le **tableau 18.1** regroupe différentes informations qui doivent permettre de classer quatre reptiles du groupe des squamates, dont un fossile. On considère le varan comme un groupe externe, sur la base d'autres études phylogénétiques.

Tableau 18.1 Caractéristiques de quatre squamates.

		python	varan	<i>Leptotyphlops</i>	<i>Pachyrhachis</i>
A	corps allongé	oui	oui	oui	oui
B	nombre de vertèbres précloacales	plus de 140	moins de 140	plus de 140	plus de 140
C	yeux	développés	développés	rudimentaires	développés
D	nombre de foramens dentaires	au plus 2	au moins 3	au plus 2	au plus 2
E	membres postérieurs	réduits ou absents	développés	réduits ou absents	développés

Les informations présumées utiles sont alors organisées en un tableau, appelé **matrice de caractères** (ou matrice taxons-caractères), dans lequel l'état ancestral d'un caractère est codé par « 0 » et l'état dérivé par « 1 ». Puisqu'un groupe externe (le varan) a été repéré, pour polariser les caractères, on retient que, pour chaque caractère, l'état présent chez le varan est une plésiomorphie (codé par 0) ; l'autre état est une apomorphie (codé par 1).

Pour faciliter la suite du travail, la **matrice de caractères** (**tableau 18.2**) peut être ordonnée suivant le nombre des occurrences des états dérivés (même si cela ne s'impose pas pour un exemple aussi simple que celui-ci).

Tableau 18.2 Matrice taxons – caractères établie à partir du tableau 18.1

	varan	<i>Pachyrhachis</i>	python	<i>Leptotyphlops</i>
A	0	0	0	0
C	0	0	0	1
E	0	0	1	1
B	0	1	1	1
D	0	1	1	1

Les caractères sont désignés par les lettres correspondant à leur ligne dans le tableau 18.1. L'état ancestral d'un caractère est codé par « 0 » ; l'état dérivé par un « 1 ».

- **Simplifier la matrice taxons-caractères**

Dans un second temps, tous les **caractères non informatifs** sont éliminés. Le caractère A (corps allongé) n'apporte aucune information ; c'est sans doute un caractère commun à tous les squamates. La ligne A n'est pas informative.

Pour les caractères B (nombre de vertèbres précloacales) et D (nombre de forams dentaires), les états se partagent de manière identique entre les taxons : ces caractères pourraient être **redondants**. Mais, compte tenu de l'indépendance apparente (anatomique, fonctionnelle et embryologique) entre les dents et les vertèbres, il est difficile d'émettre l'hypothèse que ces deux caractères aient évolué ensemble. On pourrait ainsi les retenir tous les deux dans la matrice, car plus un lien de parenté est fondé sur un nombre important de caractères, plus il paraît solide. Malgré tout, leur présence à l'état dérivé chez tous les membres du groupe (sauf le varan) fait qu'ils ne sont pas informatifs.

La matrice peut alors être simplifiée encore de ces deux lignes-là et la matrice finale ne comprend donc plus que les seules lignes C et E ([figure 18.5](#)).

- **Recenser les différentes topologies d'arbre**

Cette étape est d'autant plus complexe liaison évolutive qu'il y a un grand nombre d'organismes à classer. Elle ne pourra que rarement se faire « à la main ». Dans l'exemple que nous étudions, il y a trois animaux à classer (en dehors du groupe externe). Puisqu'il est possible de faire tourner des branches autour d'un nœud sans changer la signification de l'arbre, il y a trois topologies possibles. Elles sont présentées sur la [figure 18.5](#).

- **Confronter des scénarios évolutifs pour retenir l'arbre le plus parcimonieux**

On associe ensuite à chaque arbre les différentes étapes évolutives qui rendent compte de la distribution des états des caractères dans la matrice. Le groupe externe n'étant introduit que pour polariser les caractères, il ne porte que des caractères à l'état ancestral (état « 0 »). Chaque acquisition d'un état dérivé (« 0→1 ») correspond à un événement de transformation (par exemple une mutation) ; chaque transition vers un état ancestral (« 1→0 ») est interprétée comme un autre événement de transformation (par exemple une réversion). Pour des raisons pratiques, ce formalisme peut être représenté en utilisant une couleur spécifique par caractère.

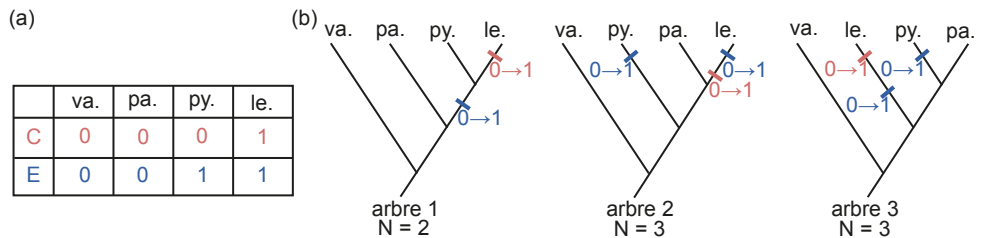


Figure 18.5 Différentes hypothèses de transformations des caractères ; (a) matrice taxons-caractères simplifiée ; (b) les trois arbres possibles et des hypothèses correspondantes.

Les animaux sont désignés par les deux premières lettres de leur nom ; la signification des caractères C et E est indiquée sur le tableau 18.2. N désigne le nombre de transitions de caractères associé à chaque arbre. D'autres scénarios pourraient être proposés sur les arbres 2 et 3, incluant des réversions. Ils ne sont pas retenus car ne conduisent pas à des arbres plus parcimonieux que les deux exemples présentés.

Parmi tous les arbres proposés, on choisit de retenir le (ou les) arbre(s) le (ou les) plus parcimonieux, c'est-à-dire présentant un nombre minimal de transformations. Dans l'exemple étudié, les arbres 2 et 3 ne sont pas retenus car moins parcimonieux. Par ailleurs, aucune information ne permettrait de retenir l'un plutôt que l'autre. L'arbre 1 est le **plus parcimonieux** mais il faudrait évidemment un peu plus de deux caractères pour le valider.

2.2 Tester l'interprétation représentée par un arbre

Un arbre phylogénétique n'est qu'une **interprétation** de la répartition des états de caractères particuliers entre différents taxons. Établir une phylogénie va au-delà de l'établissement d'un seul arbre. Différentes méthodes sont utilisées pour cela.

a) Confronter différents arbres

Une phylogénie aura d'autant plus de fiabilité qu'elle sera établie par la confrontation de plusieurs arbres produits à partir d'études fondées sur des attributs indépendants, ceux-ci pouvant être notamment des données moléculaires.

Les caractères moléculaires apportent beaucoup à l'établissement des phylogénies parce que de nombreuses molécules séquencées (ADN, ARN, protéines) sont utilisables, dès lors qu'elles sont présentes dans l'ensemble des organismes étudiés. Dans une séquence, chaque position de monomère correspond à un caractère qui peut varier par mutation ; les différents monomères sont les états de caractères. Si une même séquence n'a pas trop évolué chez différents organismes, on peut aligner chaque site un à un.

Plusieurs méthodes basées sur divers algorithmes sont ainsi utilisées pour construire des arbres. Le **tableau 18.3** en consigne trois parmi d'autres.

Tableau 18.3 Diverses méthodes de construction des arbres.

	Méthode du maximum de parcimonie	Méthode du maximum de vraisemblance	Méthode des distances <i>Neighbor-Joining</i> (NJ) (1)
Principe	Construire tous les arbres possibles.	Proposer une valeur de vraisemblance (calcul de probabilité) à un arbre donné, en prenant en compte un modèle de transformation de séquences.	Évaluer la distance séparant deux taxons en calculant le pourcentage de différences entre les nucléotides (méthode phénétique).
Choix de l'arbre	Le plus économe en transformations (principe de parcimonie).	Le plus probable (vraisemblance maximale).	Pas de choix : chaque arbre consigne les distances mesurées.
Domaine d'application	Tous les caractères (y compris moléculaires).	Caractères moléculaires.	
Avantages	Consigne l'évolution des caractères dans l'ensemble de l'arbre. Permet de distinguer convergence et homologie.		
	Rapide	Fiable	Rapide ; permet de traiter des arbres lourds.
Inconvénients	Nécessite la construction d'un nombre élevé d'arbres.	Tributaire de la qualité du modèle d'évolution. Calculs souvent très longs.	Ne permet pas de suivre l'évolution des caractères au sein de l'arbre.

(1) Cette méthode remplace une plus ancienne et moins performante, la méthode UGMA dont un exemple est présenté dans le TP 14.

La méthode du maximum de vraisemblance, considérée comme la plus fiable, se distingue notamment par le fait qu'elle ne suppose pas de savoir polariser les caractères avant d'établir l'arbre. Les plésiomorphies et apomorphies sont reconnues *a posteriori*, lors de l'analyse de l'arbre retenu. Quelle que soit la méthode de construction d'un arbre, elle doit être accompagnée de tests de fiabilité.

Voir chapitre 13, § 1.2

b) Tester la robustesse d'un arbre

La **robustesse** d'un arbre peut être définie comme son indépendance vis-à-vis du hasard de l'échantillonnage, c'est-à-dire des caractères qui ont été utilisés pour le construire.

Une des méthodes pour tester la robustesse d'une branche phylogénétique est celle du *bootstrap*. Un programme informatique recrée un grand nombre de jeux de données indépendants, par rééchantillonnage numérique à partir du jeu de données réel. Il relance ainsi la construction phylogénétique de façon itérative puis compare les résultats obtenus après un grand nombre de répétitions. Pour chaque branche de l'arbre, le programme calcule une valeur de *bootstrap* (de 0 à 100 %) traduisant le nombre de fois où cette branche a été retrouvée au fil des répétitions. Une valeur élevée (100 ou proche de 100) traduit une robustesse élevée. Les résultats de ce test sont alors consignés sur les diverses branches de l'arbre.

Remarque

Il ne faut pas confondre les valeurs de *bootstrap* portées sur un arbre et les distances mentionnées sur un arbre produit par les méthodes NJ ou UGMA ; chacune des branches de ce dernier peut porter une valeur qui correspond à la « distance génétique » qui sépare un taxon considéré de l'ancêtre commun supposé.

2.3 Utiliser les informations données par un arbre

a) Identifier un clade

Dans les différents arbres phylogénétiques des eumétazoaires qui ont pu être réalisés, le groupe des vertébrés apparaît toujours comme un **clade** ; il est donc confirmé en systématique phylogénétique. Sur le cladogramme des vertébrés (figure 18.6), les gnathostomes, les ostéichthyens, les tétrapodes, les amniotes, et les mammifères, des taxons de la systématique gradiste, sont confirmés en tant que clades dans la systématique phylogénétique. À chacun d'entre eux sont associés des synapomorphies caractéristiques (en bleu sur la figure).

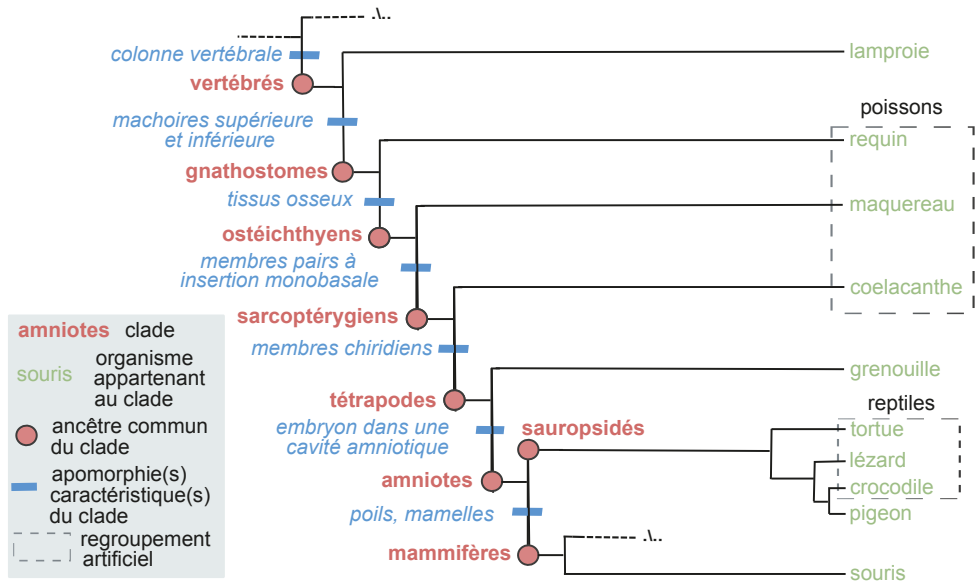


Figure 18.6 Cladogramme des vertébrés.

Une apomorphie est définie par rapport à un état ancestral du caractère considéré. Ainsi, le membre chirodien constitue un état dérivé par rapport à l'état ancestral (membre ptérygien) partagé par tous les ostéichthyens pour leurs membres pairs. Par contre, ce même état de caractère est ancestral pour tous les tétrapodes et n'est donc d'aucune utilité pour les classer entre eux (caractère non informatif dans ce cas).

En revanche, le grade des poissons apparaît comme un **groupe paraphylétique** et leur ressemblance globale relève de **symplesiomorphies** (figure 18.4b).

De même, l'ancien grade des reptiles constitue aussi un groupe paraphylétique. Par contre, le regroupement de ces reptiles et des oiseaux forme le clade des sauropsidés.

D'une manière générale, de nombreux taxons qui avaient été définis auparavant sur la base de l'absence d'un caractère, ne sont pas confirmés en cladistique. Les uns correspondent à des regroupements **paraphylétiques** (poissons, reptiles, invertébrés) ; les autres à des regroupements **polyphylétiques** comme pour les algues et les champignons qui formaient les thallophytes (chlorophylliens pour les premières, non chlorophylliens pour les seconds), sur la base d'une absence commune de tiges feuillées.

b) Valider ou non une hypothèse d'homologie

D'après le cladogramme de la figure 18.6, la colonne vertébrale, pour laquelle une hypothèse d'homologie avait pu être posée préalablement à toute étude phylogénétique, est donc une synapomorphie des vertébrés ; c'est une **homologie secondaire** (figure 18.4a).

Si une hypothèse d'homologie avait été faite pour l'aile de certains mammifères et celle des oiseaux, ce même cladogramme conduirait à l'invalider, puisque ce caractère n'est pas hérité de l'ancêtre commun des oiseaux et des mammifères. Il s'agit d'une homoplasie (figure 18.4d).

ZOOM 1

Caractères utilisés pour classer les organismes

- Le terme de **caractère** (ou attribut) désigne ce qui est observable et peut être considéré comme similaire chez deux organismes différents.

Un caractère peut être de nature morpho-anatomique (symétrie, organisation du membre), histologique ou embryologique (présence d'un tissu ou d'un feuillet embryonnaire), cellulaire (présence d'un noyau, de plastides, etc.), moléculaire (séquence nucléotidique ou peptidique) ou fonctionnel (voie métabolique, modalité de passage de la mauvaise saison, etc.)

Un caractère est **relatif** à un certain groupe d'organismes : par exemple, le caractère « organisation du membre » n'est pas utilisable chez les arthropodes qui ont des pattes et non des membres.

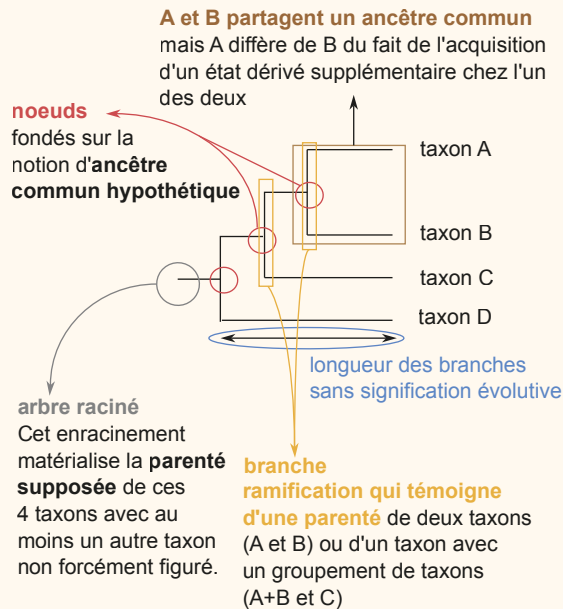
- Un caractère est présent sous **différentes formes ou états**.

La démarche cladistique repose sur la polarisation des caractères, ce qui suppose d'identifier plusieurs états d'un même caractère et de pouvoir distinguer l'état ancestral de l'état dérivé, soit avant la construction de la phylogénie (par exemple par référence à un groupe externe), soit après cette construction si elle a été réalisée en référence à un modèle d'évolution. Il ne faut donc pas confondre le caractère « organisation du membre », avec ses états « membre ptérygien » (nageoire paire d'un poisson) et « membre chiridien » (membres des vertébrés terrestres). Un état dérivé résulte de la transformation d'un état ancestral, ce qui est parfois qualifié d'innovation, c'est-à-dire un changement héritable, sans que ce terme soit synonyme de progrès.

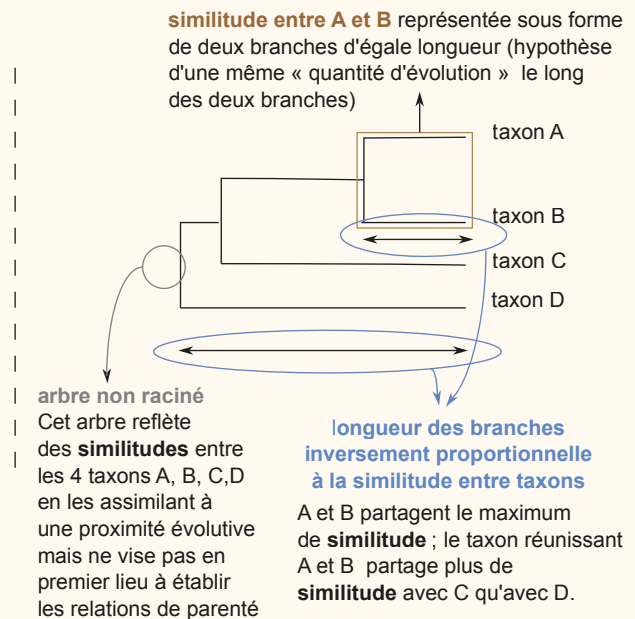
ZOOM 2

Phénogramme ou cladogramme

Qu'il s'agisse de la systématique cladistique ou de la systématique phénétique, les apparentements sont représentés par la production d'arbres (ou dendrogrammes) qui, malgré leur ressemblance, diffèrent aussi bien par leur méthode de construction que par leur interprétation. Un phénogramme représente le **degré de similitude globale** entre des taxons, mais n'apporte rien sur la nature de leur ressemblance (homologie secondaire, homoplasie ou symplesiomorphie) à la différence d'un cladogramme dont la lecture permet justement de préciser la **nature des ressemblances**.



CLADOGRAMME



PHÉNOGRAMME

Organisations comparées d'un cladogramme et d'un phénogramme.

DÉCOUVERTE 1

Taxonomie et classification : un bref historique

L'objectif des premiers systèmes de classement était d'organiser la biodiversité, selon des cadres de pensées créationnistes, desquels la notion d'évolution était totalement absente (fixisme).

- Karl von Linné (1758, *Systema naturae*) : **dénomination binominale** des espèces (nom de genre suivi d'un terme désignant l'espèce) et classification hiérarchique en **sept rangs**. Il fonda la taxonomie moderne. De nombreux travaux ont ensuite abouti à des classements dans lesquels les mêmes rangs étaient retrouvés, ce qui semblait, selon les idées de l'époque, traduire l'existence d'un ordre d'inspiration divine.

- B. et A.L de Jussieu : principe de « **subordination des caractères** »
Parmi les différents caractères recensés, sont d'abord distingués les « caractères dominants » (retrouvés dans un nombre important de taxons) ; puis sont pris en considération les autres caractères, qui leur sont donc « subordonnés ». La présence de vertèbres (vertébrés) est un caractère de rang supérieur à celui de la présence de poils et de glandes mammaires (mammifères) ; de même pour la présence de tissus conducteurs (trachéophytes) et celle d'un ovule protégé dans un fruit (angiospermes). Ainsi un taxon d'ordre supérieur (ordre, famille...) est défini par des caractères placés plus haut dans la hiérarchie que ceux d'un taxon inférieur (genre, espèce).
- G. Cuvier : **classification des animaux** (1817)
Il reconnaît quatre embranchements fondamentaux, auxquels il attribue un plan d'organisation particulier, sans aucun terme de passage entre ces embranchements.
À la fin du XVIII^e siècle, différents naturalistes remettent en cause la fixité des espèces et envisagent simultanément la possibilité qu'elles se modifient.
- J.B. Lamarck : **première grande théorie transformiste** (*Philosophie zoologique*, 1809)
Il énonce 3 principes : les espèces se modifient au cours du temps et se transforment en d'autres espèces ; la nature montre une tendance à la complexification et au progrès ; l'hérédité des caractères acquis sous la contrainte du milieu, conduit à la transformation des espèces.
- E. Geoffroy Saint Hilaire : « **loi d'unité de composition organique** » (1828)
S'opposant à G. Cuvier, il défend l'idée d'une unité fondamentale du règne animal ; tous les animaux sont bâtis sur le même plan, leur diversité se résume à des caractères discrets.
- C. Darwin : idée de **descendance avec modification** (*L'origine des espèces* 1859)
Avec le naturaliste britannique A.R. Wallace, au lieu d'invoquer des causes externes qui orientent les transformations des espèces, il émet l'idée de variations internes et de sélection naturelle.
L'acceptation de ces idées évolutionnistes ouvre la voie à une classification des êtres vivants en fonction de leur apparentement. Elle permet de passer d'un système hiérarchique de description du vivant à une véritable classification.
- E. Haeckel : **première classification des êtres vivants** (1866)
Il qualifie de phylogénie sa classification qui prend la forme d'un **arbre** inspiré des arbres généalogiques. Selon lui, la hiérarchie des caractères peut être déduite de l'étude du développement (ontogenèse) des organismes : durant l'ontogenèse d'un organisme, les caractères dominants apparaissent avant les caractères subordonnés. Une formule résume son travail phylogénétique : « l'ontogenèse récapitule la phylogenèse ». L'arbre basé sur la hiérarchie de la distribution des caractères devient alors une représentation des liens de parenté entre les organismes vivants. Mais la phylogénie de E. Haeckel repose sur une conception lamarckienne de l'évolution, puisqu'il assimile arbitrairement le principe de subordination des caractères à une chronologie évolutive tournée vers une complexification croissante et progressiste des organismes.
- W. Hennig : **systématique phylogénétique** (1950)
Il a posé les bases de la phylogénie actuelle, en proposant l'idée de regrouper les organismes selon des caractères hérités d'un ancêtre commun (ensemble hypothétique d'états dérivés de caractères).

Réviser

PROBLÈME : Comment mener une étude phylogénique sur un groupe constitué de différents êtres vivants et comprenant éventuellement des fossiles ?

1 DES FAITS

RECENSER différents caractères disponibles pour classer les membres du groupe.

2 UNE METHODE :

TRIER les caractères utiles pour classer : retenir ceux existant au moins sous deux états et permettant de formuler des hypothèse d'**apparentement**.

POLARISER les caractères : distinguer caractères primitifs/caractères dérivés ; choisir un groupe externe.

IDENTIFIER les ressemblances suggérant une parenté ; **homologies primaires = hypothèses d'apparentement**.

CONSTRUIRE et EXPLOITER un **arbre** :

- organiser une **matrice** de caractères
- retenir l'arbre le plus **parcimonieux**
- tester la robustesse de l'arbre

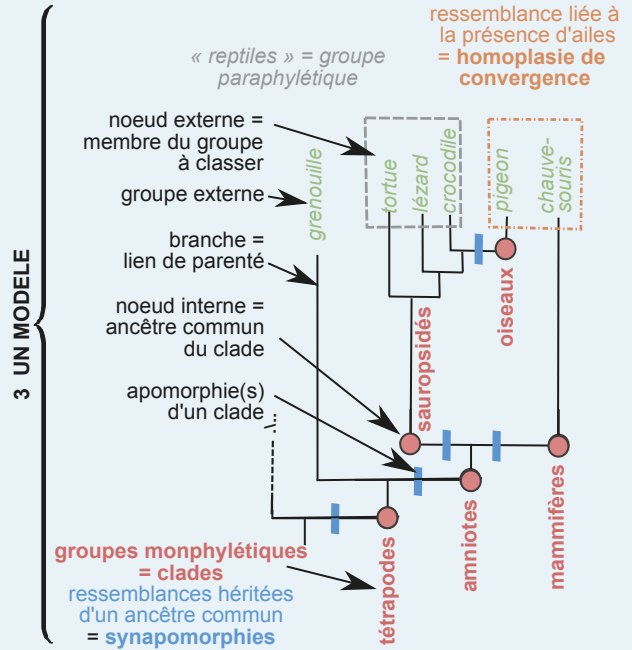


Figure de synthèse L'approche phylogénétique cladistique.

S'entraîner

QCM de connaissances

- 1 Repérez la définition d'une synapomorphie.
 - a. Un caractère partagé par un groupe d'êtres vivants.
 - b. Un caractère ancestral partagé par un groupe d'êtres vivants.
 - c. Une homologie primaire partagée par un groupe d'êtres vivants.
 - d. Une homologie secondaire partagée par un groupe d'êtres vivants.
- 2 À propos des arbres phylogénétiques.
 - a. Ils se construisent uniquement par la démarche phénétique.
 - b. Les taxons se placent sur les nœuds internes.
 - c. Ils permettent d'identifier *a posteriori* les homologies secondaires.
 - d. Il faut obligatoirement y placer une racine.
 - e. Ce sont des interprétations des relations d'apparentement entre taxons.
 - f. Ils sont plus fiables quand ils sont établis sur des données moléculaires.

QCM à partir de documents

1 Le tableau 18.4 recense les états de 4 caractères pour 6 vertébrés.

Tableau 18.4 Matrice taxons - caractères.

	Écailles épidermiques recouvrant...	Fenêtres préorbitaires	Gésier	Plumes
aigle	les pattes	présentes	présent	présentes
autruche	les pattes	présentes	présent	présentes
boa	tout le corps	absentes	absent	absentes
crocodile	tout le corps	présentes	présent	absentes
grenouille	écailles absentes	absentes	absent	absentes
lézard	tout le corps	absentes	absent	absentes

- a. La grenouille est le groupe externe.
 - b. L'état « écailles épidermiques recouvrant tout le corps » est dérivé.
 - c. On peut supprimer l'un des caractères « fenêtre orbitaire » ou « gésier » car ils sont redondants.
 - d. Les plumes sont une synapomorphie de l'autruche et de l'aigle.
 - e. L'absence de fenêtre orbitaire est une synapomorphie de la grenouille, du boa et du lézard.
- 2 La figure 18.7 présente deux arbres regroupant les 6 vertébrés de la question 1. Chacun est représenté par l'initiale de son nom.
- a. L'arbre A correspond à celui qui peut être établi à partir du tableau 18.4.
 - b. Crocodile, aigle et autruche forment un groupe monophylétique.
 - c. Grenouille, lézard et boa forment un groupe monophylétique.
 - d. L'arbre reste correct si on intervertit l'aigle et l'autruche.
 - e. L'arbre reste correct si on intervertit l'aigle et le crocodile.

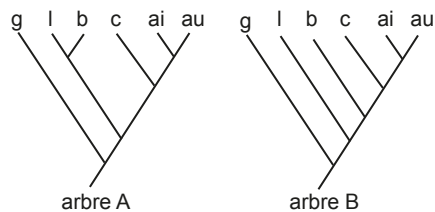


Figure 18.7 Arbres phylogénétiques avec 6 vertébrés.

Question de synthèse courte

Les principes de la classification du vivant.

Reconstruction phylogénétique

Activités pratiques

PLAN DU CHAPITRE

- 1 Construire et utiliser un cladogramme
- 2 Construire un phénogramme

INTRODUCTION

Les hypothèses sur les liens de parenté entre les êtres vivants sont matérialisées par des arbres phylogénétiques qui peuvent être établis par différentes méthodes. Elles permettent de reconstruire des phylogénies plus cohérentes que les classifications antérieures. Deux de ces méthodes seront illustrées dans ce TP.

La méthode du maximum de parcimonie (présentée au chapitre 18, § 2.1) sera mise en œuvre à partir de caractères morpho-anatomiques d'un groupe d'oursins. Une méthode phénétique simple permettra ensuite d'exploiter des données moléculaires.

1 Construire et utiliser un cladogramme

Les étapes du raisonnement cladistique seront illustrées à partir de l'étude phylogénétique d'un groupe d'oursins (Echinoidea, Echniodermata) qui offre de nombreux caractères morphologiques utilisables pour l'analyse phylogénétique, particulièrement au niveau du squelette ou test. Le plan d'organisation de ce groupe n'est pas indispensable à connaître pour effectuer le raisonnement phylogénétique. La [figure TP14.1](#) présente les principaux caractères recensés lors de l'analyse cladistique. De nombreuses autres informations et photographies sont disponibles sur le site <http://www.snv.jussieu.fr/bmedia/aimar2/PAGES/25.htm>

1.1 Recenser les caractères utilisables pour établir une phylogénie

L'exosquelette d'un oursin, ou **test**, est constitué d'une enveloppe calcaire. Les oursins dont le test hémisphérique ou elliptique, présente une symétrie radiale, sont dit réguliers ([figure TP14.1a et b](#)) ; ceux dont le test présente une symétrie bilatérale sont dits irréguliers et ont un mode de vie fouisseur ([figure TP14.1c et d](#)).

À partir de caractères morphologiques, nous allons chercher à émettre une hypothèse sur les relations phylogénétiques entre les quatre taxons d'oursins suivants :

- les camarodontes (représentant choisi : *Paracentrotus lividus*) ;
- les cidaroides (représentant choisi : *Cidaris cidaris*) ;
- les clypeasteroides (représentant : *Clypeaster subdepressus*) ;
- les spatangoïdes (représentant choisi : *Echinocardium mediterraneum*).

Le choix d'un unique représentant par groupe suppose que la monophylie de chacun des quatre taxons ait déjà été établie par des travaux antérieurs. La polarisation des états des différents caractères se fera par comparaison avec un représentant d'un autre échinoderme considéré comme groupe externe, une étoile de mer (astéroïdes).

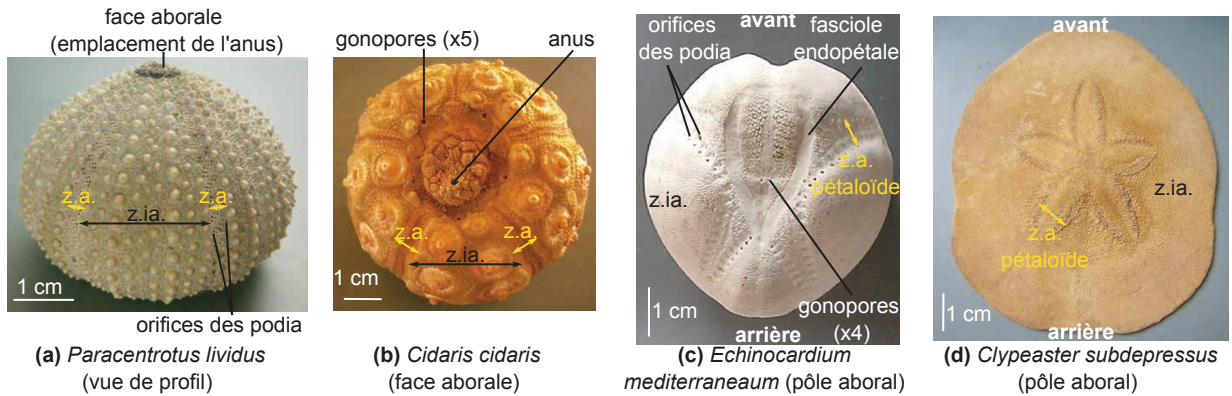


Figure TP14.1 Quelques caractéristiques des tests des oursins.

(Clichés C. Aimar, Biologie et Multimedia, Sorbonne Université <https://rnbio.sorbonne-universite.fr>).

Le **test** est formé de plaques calcaires, assemblées en **zones ambulacraires** (z.a) et **interambulacraires** (z.ia). Diverses structures y sont localisées : **podia** (expansions rétractibles sortant du test au niveau des pores de la zone ambulacraire) et **radioles** (= piquants) sont les plus visibles sur l'animal vivant.

La position de la bouche définit le **pôle oral** du test ; à l'opposé se trouve le **pôle aboral**. Chez de nombreux oursins, le seul élément visible au niveau de la bouche est la **lanterne d'Aristote** ; elle est formée de pièces calcaires articulées et de leurs muscles qui permettent d'assurer les mouvements de cinq dents (cette structure n'est pas visible sur les photographies ci-dessus, compte tenu de leur orientation). L'émission des gamètes s'effectue par les pores génitaux ou **gonopores**. Sur certaines espèces (**c**) des bandes étroites de localisation précise portent des radioles plus fins ; elles constituent les **fascioles** qui permettent l'identification des espèces.

Un recensement des caractères morphologiques des oursins a conduit à établir la matrice taxons-caractères de la **figure TP14.2**.

caractère	1	2	3	4	5	6	7	8	9	10	11	12
astéroïdes	a	?	b?	a	b	a	?	?	b?	b?	b?	b
cidaroides	a	a	b	a	b	a	a	a	b	b	b	a
camarodontes	a	a	b	a	b	b	b	a	b	b	b	a
clypeastéroïdes	b	b	b	b	a	b	c	b	a	a	b	a
spatangoïdes	b	a	a	b	a	b	c	b	a	a	a	b

- | | |
|---|---|
| <p>1. Symétrie a : pentaradiée
b : bilatérale</p> <p>2. Forme générale a : ovoïde
b : plat</p> <p>3. Forme particulière a : en cœur,
b : non</p> <p>4. Position de l'anus a : centre de la face aborale
b : ailleurs</p> <p>5. Nombre de gonopores a : 4
b : 5</p> <p>6. Plaques ambulacraires a : simples, libres,
b : soudées</p> | <p>7. Tubercules des plaques interambulacraires a : un seul
b : plusieurs non identiques
c : plusieurs identiques</p> <p>8. Morphologie des piquants a : longs et épais
b : courts et fins</p> <p>9. Aires ambulacraires de la face orale a : forme pétaloïde
b : non</p> <p>10. Rôle respiratoire des podia de la face orale a : oui
b : non</p> <p>11. Fascioles sur le test a : présents
b : absents</p> <p>12. Lanterne d'Aristote a : présente
b : absente</p> |
|---|---|

Figure TP14.2 Matrice taxons-caractères.

À chacun des douze caractères (en rouge) a été attribué un numéro (en gras) ; les différents états des caractères sont codés par des lettres (a, b, ou c).

1.2 Éliminer les caractères non informatifs

La polarisation d'un caractère n'est ici possible que lorsque son état est connu pour le groupe externe (astéroïdes). On réalise assez rapidement que ce n'est le cas que pour les caractères 1, 4, 5, 6 et 12. Les autres ne sont pas informatifs, car **non polarisables**. De toute façon, les caractères 3 et 11, même si l'état b était confirmé pour le groupe externe, n'auraient pas eu non plus de valeur informative car l'état a (dérivé) n'aurait eu **qu'une seule occurrence**.

On peut noter une **redondance** de la répartition des états entre les caractères 1 et 4 : or la symétrie bilatérale est justement due à la position excentrée de l'anus. On ne conservera donc qu'un seul de ces deux caractères (par exemple, le caractère 1). Il y a aussi une redondance de la répartition des états des caractères 4 et 5. L'absence d'une gonade (donc d'un gonopore) pourrait être une conséquence de la migration de l'anus loin du pôle aboral, mais dans ce cas, éliminer un des deux caractères revient à faire une hypothèse *a priori* sur les modalités d'évolution. On conservera donc les deux caractères.

Au final, l'analyse cladistique sera réalisée sur les **quatre caractères 1, 5, 6 et 12**. Le **tableau TP14.1** résume pour chacun la répartition de l'état ancestral (codé 0) ou dérivé (codé 1).

Tableau TP14.1
matrice des caractères simplifiée et ordonnée.

caractère	1	5	6	12	occurrences (état 1)
astéroïde	0	0	0	0	0
cidaroïdes	0	0	0	1	1
camarodontes	0	0	1	1	2
spatangoïdes	1	1	1	0	3
clypeasteroïdes	1	1	1	1	4
occurrences (état 1)	2	2	3	3	

1.3 Procéder à l'analyse cladistique

Comme on ne change pas la signification d'un arbre en faisant tourner des branches autour d'un nœud, on peut recenser douze topologies possibles d'arbre comme le type 1 de la **figure TP14.3a** et trois du type 2, soit quinze arbres possibles au total.

L'analyse cladistique consiste alors à placer, pour chaque topologie, les hypothèses de transformations de caractères. La **figure TP14.3b** donne deux résultats de cette étape dont l'arbre le plus parcimonieux qui correspond à 5 transformations (5 pas).

Voir chapitre 18, § 2.1c

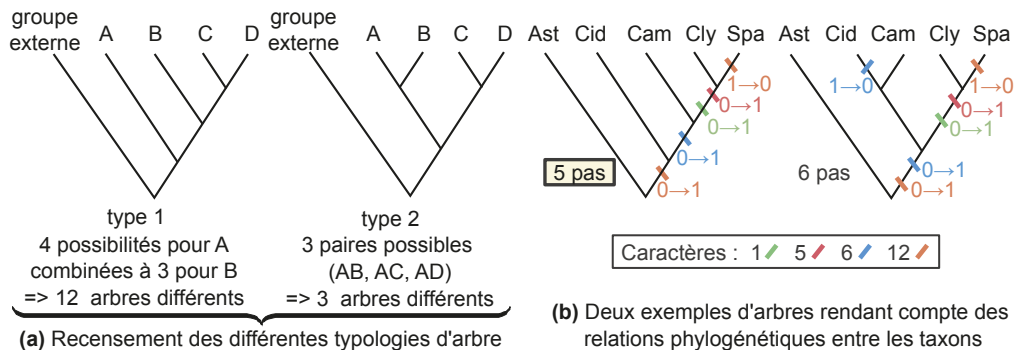


Figure TP14.3 Analyse cladistique à partir du tableau TP14.1.

1.4 Utiliser le cladogramme

On peut utiliser cet arbre pour revenir sur l'évolution des caractères eux-mêmes (**figure TP14.4**). La soudure des plaques ambulacraires (caractère 6) apparaît comme une **synapomorphie** du clade (camarodontes, clypeastéroïdes, spatangoïdes). Les oursins irréguliers (clypeastéroïdes, spatangoïdes) apparaissent comme un clade dont la symétrie bilatérale (caractère 1) et le nombre réduit

de gonopores (caractère 5) sont deux synapomorphies. Enfin, la réversion du caractère 12 correspond à la perte de la lanterne d'Aristote chez les spatangoides. Celle-ci a été déduite de l'analyse phylogénétique et non comme une hypothèse supposée *a priori* (si les spatangoides étaient situés sur le nœud profond de l'arbre on aurait très bien pu obtenir une absence primitive !).

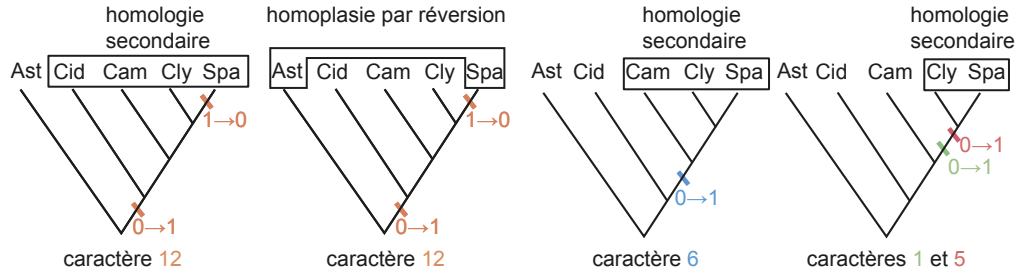


Figure TP14.4 Analyse cladistique à partir du tableau TP14.1.

2 Construire un phénogramme

On cherche à établir un arbre phylogénétique avec des plastes d'organismes de la lignée verte (*Arabidopsis thaliana*, *Cyanophora paradoxa*, *Porphyra purpurea*) et une cyanobactérie (*Anabæna sp.*). Pour cela, on compare des séquences partielles des gènes d'ARNr 16S du stroma des plastes ou du cytoplasme des cyanobactéries. La figure TP14.5a indique la distance entre chaque couple d'espèces, définie comme le pourcentage de nucléotides qui diffèrent entre les séquences des espèces considérées.

On utilise la méthode UPGMA (*Unweighted Pair Group Method using Averages*, figure TP14.5), qui n'est plus guère utilisée que pour les exercices, car elle a l'inconvénient majeur de ne pas tenir compte des taux de mutations différents sur les différentes branches. Elle a, en revanche l'avantage d'être facile à mettre en œuvre sans moyens informatiques. Les scientifiques lui préfèrent la méthode du *Neighbor-Joining*.

Voir chapitre 18, tableau 18.3

	<i>Arabidopsis</i>	<i>Anabæna</i>	<i>Cyanophora</i>	<i>Porphyra</i>
<i>Arabidopsis</i>	0			
<i>Anabæna</i>	38	0		
<i>Cyanophora</i>	20	29	0	
<i>Porphyra</i>	12	31	17	0

(a) Matrice des distances

% de différences de séquence des ARN 16 S

3. Créer un nouvel ensemble U incluant i et j.

4. Si i et j sont les deux derniers ensembles, alors l'arbre est complet.

Si non, recalculer une matrice de distances en prenant comme distance

de U à chaque autre taxon k, la moyenne des distances d_{ki} et d_{kj} $d_{Uk} = (d_{ki} + d_{kj})/2$.

	U	<i>Anabæna</i>	<i>Cyanophora</i>
U	0		
<i>Anabæna</i>	$(38+31)/2 = 34,5$	0	
<i>Cyanophora</i>	$(20+17)/2 = 18,5$	29	0

(b) % de différences entre U et les autres taxons

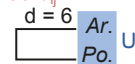
	U'	<i>Anabæna</i>
U'	0	
<i>Anabæna</i>	$(34,5+29)/2 = 31,75$	0

(c) % de différences entre U' et la cyanobactérie

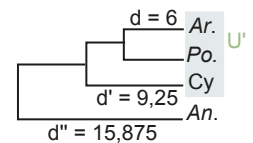
1. Trouver les taxons i et j pour lesquels la distance d_{ij} est la plus petite.

○ *Arabidopsis* et *Porphyra*

2. Mettre la racine à égale distance de i et de j, soit à $d_{ij}/2$.



5. À partir de cette nouvelle matrice, recommencez à l'étape 1.



(d) Phénogramme

Figure TP14.5 Mise en œuvre de la méthode UPGMA à partir de séquences d'ARN 16S.

Le groupe U' est constitué par les plastes des organismes eucaryotes ; la cyanobactérie correspond au taxon le plus éloigné des 3 autres, ce qui apparaît comme un résultat cohérent.