

**Zeitschrift:** Archives des sciences et compte rendu des séances de la Société  
**Herausgeber:** Société de Physique et d'Histoire Naturelle de Genève  
**Band:** 49 (1996)  
**Heft:** 2: Archives des Sciences

**Artikel:** Roots and xylem of phylogenetics  
**Autor:** Bittar, Gabriel  
**DOI:** <https://doi.org/10.5169/seals-740419>

### **Nutzungsbedingungen**

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. [Mehr erfahren](#)

### **Conditions d'utilisation**

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. [En savoir plus](#)

### **Terms of use**

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. [Find out more](#)

**Download PDF:** 02.05.2026

**ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>**

**Communication présentée à la séance du 8 février 1996**

## ROOTS AND XYLEM OF PHYLOGENETICS

BY

**Gabriel BITTAR\***

### ABSTRACT

**Roots and xylem of phylogenetics.** - Historical roots of phylogenetics are briefly presented, from Antiquity to the beginning of the 20th century. Then the main corpus of basic technical concepts and methods of this deeply rooted yet strongly growing science is succinctly described.

### **The historical foundations of phylogenetics**

Mankind has always tried to put order in the apparent chaos of the natural world, by using more or less sophisticated and appropriate methods. In the 18th century, both the development of classificatory methods, and the accumulation of naturalist knowledge through observation, had allowed the development of the science of *systematics*. Its main aim was to help naturalists in their task of proper identification of living beings. In the 19th century, progress in *palaeontology* (from the Greek *palaios* = ancient, and *ôn, ontos* = being, existence) led scientists to propose a more fundamental approach of systematics, one that would put order in both present and past living beings, in evolutionary terms. Thence was defined the science of *phylogenetics* (Gr. *phylon* = tribe, and *genos* = birth, origin, *gennêtikos* = proper to generation), which objective is to establish a phylogeny of past and present living beings, i.e. to appraise systematically the relationships between them by tracing family links and biological descent : this task can be considered as analogous to the establishment of a *genealogical tree*, with the difference that it is not individuals that appear in the tree, but biological groups, or *taxa* (a modern term derived from the word taxonomy introduced, in 1813, by A.-P. DE CANDOLLE, the great botanist from Genève: Gr. *nomos* = law, and *taxis, taxéôs* = order, arrangement).

Simply said, systematics and its modern offshoot, taxonomy, are a necessary preliminary to the investigation of the causes involved in the resemblance and dissimilarities found in living species, this investigation being done under the phylogenetic model which is a most direct way for a correct evaluation of the evolutionary events having

---

\* G. Bittar, ch. du Pont-Noir 9A, CH-1226 Thônex.

caused these (dis)similarities. This may seem to be rather an obvious objective *per se*, even if not trivial to accomplish. Yet what seems patently necessary nowadays was not at all that evident only two centuries ago. So let us briefly get an epistemological remembrance of the historical roots of phylogenetics.

At the end of the 18th century (“le Siècle des Lumières”), an important transformation had occurred in the philosophical conceptualisation of the natural world. It was the end of the *typological* view of the different components of nature, a view that was deeply rooted in occidental thinking since the establishment of PLATO's *Akadêmia* and ARISTOTLE's *Lykeion*, in the city of the Athenians, in, respectively, 385 and 335 BC.

Better than any philosophical digression, a classical example illustrates what was at stake here. Fire for long had been considered as a fundamental *element*, i.e. a fundamentally unchangeable component *animating*, in a way, the events that may happen in the world. Now and henceforth, it would be considered as a *physical* phenomenon, combustion, that could be understood in chemical ways: thus fire is the *result* of combustion, not the *cause* of combustion. This example illustrates how, since the 18th century, the rationalist-materialist-reductionist-mechanistic-experimentalist view of nature has predominated in informed circles, and has allowed an extraordinary development of scientific knowledge.

The antique roots of the *rationalist* view can be traced back to great thinkers who lived in a city of the Aegean-Anatolian coast of archaic Greece, to the minds of the Milesian “physicist”-philosophers: successively, from the beginning to the end of the 6th century BC, THALES, ANAXIMANDER and ANAXIMENES. After some detours, at the end of the 5th century BC, in the middle of the Greek classical period, the “atomists” of the northern-Aegean city of Abdera, particularly the great rationalist encyclopedist DEMOCRITUS (who was interested, among other things, in botany and zoology), had developed further this view into the direction of *materialism* and *mechanistic reductionism* (it must be noted that, generally, the Greek materialist *philosophers* were not atheists, in the same way as those in favour of a strict separation of State and Church are not necessarily anti-clerical). All these Greek philosophers (Gr. philos = friend, sophos = sage through knowledge) paved the way to the next great encyclopedist, ARISTOTLE of Stagira (another northern-Aegean city).

From the middle of the 4th century to his death in 322 BC, Aristotle recentered *philosophia* towards *epistêmê* (science, knowledge) and “*physiologia*”, i.e. away from the abstract geometrical-metaphysical speculations of PLATO and towards the observation of human society and nature. It can be noted that, even if Plato's proposals in zoological classification were very superficial, other people in the Academy, such as SPEUSIPPUS, his nephew and successor (from 347 to 339), were also discussing these problems, with a methodological approach that was much more scientific.

On the other hand, even though Aristotle rejected Plato's rather anti-scientific position and doctrine of the existence of a shadowy world of pure ideas and forms, the sole that was considered by the founder of the Academy as worth of (geometrical) “study”, his own view of the world was so heavily influenced by his master's *ontolo-*

*gical* notion of ideal form and *teleological* notion of final cause, that he fundamentally remained an elementarist and a *finalist*, unwilling to adopt the mechanistic and materialist view of the “atomists” (a view of nature that would prove, two millenaries later, to be one of the prerequisites for more fundamental developments in science).

In life sciences, ARISTOTLE'S own way of compromise between the materialism of the “physicists”-“atomists” on the one hand, the “Academic” idealism of Plato on the other hand, would lead him to propose a *categorisation* (Gr. katêgoria = statement, quality attributed to an object) scheme of living beings which, though founded on the notion that these should be considered under a *plurality of characters (differentiae)*, a notion that is the basis of modern systematics, was also founded on a *typological* hierarchy of classification (Gr. typos = impressing of mark). Hence he kept ANAXIMANDER'S notion of archetypes (Gr. archaios = ancient), adopted by PLATO, in which an ancient type could be considered as an ideal type, with, corresponding to each natural form, an ideal prototype (Gr. prôtos = first, primitive). Aristotle's scientific thinking was also profoundly influenced by his belief, which he shared with Plato, in the superiority of a class-bound, anti-democratic society, and accordingly his categorisation scheme was first and foremost a *hierarchization* scale from the most imperfect to the most perfect, with mankind as the supreme and superior animal category (and gods and noble male humans at the highest positions...).

The important logical consequence of Aristotle's views was that if there ever was any kind of transformation of living species (as ANAXIMANDER, and EMPEDOCLES of Acragas in the middle of the 5th century, had had the intuition), this could only happen, as Plato had stated, through their *degradation*, in the general direction of decadence and decay. Therefore, because of his adherence to finalism, ARISTOTLE, despite his sharp intelligence and encyclopædic knowledge, was unable to integrate the powerful notion of randomness that the genius of CHARLES DARWIN would propose 22 centuries later, and thus unable to conceptualise natural evolution.

Consequently, ARISTOTLE could not in any way be considered as the father of phylogenetics. Despite this, as already said, he is the historic father of systematics, because of his insistence, even within a typological framework, on the necessary factual observation of the many characters of living beings, free of the imaginary, the mythical, the sacred, the magical. And indeed Aristotle was a great observer, a real naturalist who can be considered (among other achievements) as the first zoologist, and it is quite understandable that Charles Darwin, himself a great observer, admired him profoundly and considered him as possibly the greatest observer in history.

For the lasting glory of Greek science, the successor of Aristotle at the head of the Lyceum, from 323 to 288/7 BC, Tyrtamos of Eresos (whom the peripatetic master had renamed THEOPHRASTUS), was also a great encyclopedist. Unlike Aristotle, he had no metaphysician tendencies at all, but was a determined disciple and, generally speaking, he accepted the philosophical framework established by the peripatetic master, including the doctrine of the final causes. But he also was a determined empiricist, even, to some extent, an experimenter, and his great ability to observation led him to reco-

gnise many natural manifestations that did not speak so much in favour of this teleological doctrine. He pointed out many aspects of nature indicating that, indeed, many things happen at random and without end. To illustrate the scientific capabilities of Theophrastus, suffice it to say that he pointedly remarked that fire being able to generate itself and requiring a material substratum, "it seems absurd to call this a primary element and as it were a principle, if it cannot exist without matter". Though most of his written works have not reached to us, this cautious and meticulous scientist is undeniably the father of petrology and mineralogy, as well as of botany. Hierarchical prejudgements not being the rule in botany (contrary to zoology), he was able to develop another classification approach devoid of the Aristotelian notion of zoological series of perfection, even stating that nature does not necessarily... "possess necessity". Yet, because of his caution as a theorist, despite of all his intellectual qualities, Theophrastus was unable to develop on the theoretical aspects of Aristotle, in fact adhering reverently to these, with some arrangements : thus, even though the universe could not really be considered as perfectly well-ordered, at least could it be considered as generally well-ordered...

Greek science would continue to develop during the three centuries following Aristotle's death (and the death of ALEXANDER THE GREAT, in 323 BC), known as the Hellenistic age. It is during this age, at the beginning of the 3rd century BC, that were founded in Alexandria of Egypt, by the first two Greek kings (PTOLEMY I AND II) of the Lagid dynasty now reigning on the multi-millenary country of the Nile, these two prestigious institutions that were the Library and, some years later, the *Mouseïon* (other Museums would be founded elsewhere, and another renowned Library would be founded in the north-western Anatolian city of Pergamum, under the Greek dynasty of the Attalids).

Alas, despite this generally favourable background for sciences, life sciences (except medicine) were not considered as immediately or politically useful, or prestigious, and were left on the side-track, even in this state-funded centre entirely devoted to research that was the Museum. Consequently, even in the realm of the life sciences, where it is the least tenable, the finalist Aristotelian view would be able to hold on up to the 19th century !

The Hellenistic society was entirely devoted to *status quo*, and considered change as synonymous to degeneration, accordingly immoral and dangerous, and definitely uninteresting as a subject of study. No wonder hence that biological sciences, except medicine, could not develop much further. Nevertheless, it is a pity that the Romans were rather impervious to philosophy and to any science that wasn't military or, *à la rigueur*, strictly clinical medicine, with the result that the three centuries of the pagan Roman Empire were rather an a-philosophical and a-scientific desert, relatively to the incredible intellectual inventivity that prevailed, in the eastern Mediterranean regions of Hellenic culture, from the 6th to the 2nd centuries BC.

The rebirth of Platonism in the 3rd century AD did not help natural sciences. This turbulent century was a turning point for philosophy and science, because from about

200 AD, in a slowly declining Empire where the general mood was of pessimism, the main effort would, more and more, tend to be spent on preserving knowledge through simplifying commentaries, rather than on increasing it. Despite the efforts of some individuals, things became even worse in the 4th century, when the Empire, after the rapid growth in wealth and power of the Christian Church since the middle of the 3rd century, officially became Christian, and subsequently, considering reason and knowledge as contrary to revelation and faith, also became, for political and ideological reasons, anti-philosophy and anti-science.

But the Platonist ideal typology and the Aristotelian concept of categorisation were generally agreed by the early Christian church, allowing the written works of these philosophers to escape the great book burning that was the fate for most other Greek philosophers' and scientists' works during the 4th century. The end result of this was that a number of PLATO'S and ARISTOTLE'S writings survived in the Byzantine Empire, and, in the 12th century, could be successfully reintroduced by the Arabs in Roman Catholic Europe, where they would become considered as somewhat scholastically sacred. This consecration did not really help the occidental philosopher-scientists inspired by the 13th century experimentalist PIERRE "PEREGRINUS" DE MARICOURT, his admirer, the clergy-persecuted English philosopher ROGER BACON (1214-1294), and his brilliant compatriot and spiritual heir (also clergy-persecuted), GUILLAUME D'OCKHAM (c.1285 - c.1349). Nor did it help the European post-Renaissance naturalists who, taking advantage of GUTENBERG'S one-century old wonderful invention of typographic printing, tried, from the middle of the 16th century to the end of the 18th century, to put order and classification in the great variety of living forms that the European explorers were describing or bringing from all around the world.

Despite an intellectually and religiously unfavourable environment, these naturalists were able to make an extraordinary pioneering job, that is too often forgotten in the history of science. Let us simply mention, in the 16th century, the zoologists KONRAD GESNER, EDWARD WOTTON, GUILLAUME RONDELET and PIERRE BELON, the botanists ANDREA CESALPINO (the "first real systematist" according to Linnæus) and JOHANN BAUHIN; in the 17th century, the naturalists GASPARD BAUHIN (brother of the preceding) and JOHN RAY, the botanists PIERRE MAGNOL and Joseph Pitton de TOURNEFORT; in the 18th century, the naturalists LINNÆUS and MICHEL ADANSON (founder of numerical systematics), the zoologists VITALIO DONATI, BUFFON and CUVIER, and three botanist members of the JUSSIEU family. In their description and classification task, together with physiologists and geologists, they paved (not always willingly...) the scientific way to the abandonment, at the end of the Enlightenment century, of the typological view of nature : living beings would no more be considered as archetypes or prototypes (more or less degenerate) of the creation or of the origins, but as *biological* organisms that may be grouped together, or distinguished between, strictly to their *observable* attributes. Therefore, scientists would no more speak of the (intrinsic or ideal) character of an animal or a plant, but of the distinctive characters defining their biological characteristics.

Thus became possible the concept of the *transformation* of natural forms of life, that the naturalist Jean-Baptiste Monnet de LAMARCK (1744 - 1829), an adept of Ockham's nominalism, propounded. By showing, with a lot of proofs, that natural history showed a propensity to increasing complexity of life forms, thus turning down on its head the antique concept of their ineluctable degradation, he definitely shattered the Platonist-Aristotelian concept of a hierarchized scale of life, in favour of the much more naturalistic concept of a tree of life, with ancestral simple organisms at the root and more complex ones at the top, the links being thus organised in a genealogical rather than reticular way. Being contrary to more than two millenaries old habits, and to deeply entrenched religious thinking, Lamarck's notions took many decades to be accepted. Many enlightened people contributed, during the first half of the 19th century, to acceptance of transformism, among them the naturalist Karl Friedrich von KIELMEYER, the zoologists GEOFFROY-SAINT-HILAIRE father and son, or the geologist CHARLES LYELL.

The next step was to provide a mechanism for transformation, that could explain the following apparent paradox, resulting from biological observation of life on one hand, paleontological discoveries of fossils on the other hand : how could there be both stability of the life forms in the short term, and transformation in the long term ? There already was, as a first element of solution, the structure of the phylogenetic tree. The second element to the solution to this riddle came from the extraordinary work of the naturalist CHARLES DARWIN (1809 - 1882), "On the origin of species by means of natural selection", published in 1859 and based on an enormous corpus of observations. In this fundamental work, the thesis of which had been confirmed by the own observations of the naturalist and zoogeographer ALFRED RUSSEL WALLACE, it was proposed that the modification of species within the long term was the result of the *random* (and generally rare, hence the general stability of species in the short term) appearance of morphological mutations within a population, and of the *natural selection* (positive or negative) of the individuals (mutated or not) within it. In this powerfully simple concept, completely materialist and mechanistic, *evolution* (from now on this word, proposed by Lyell, would replace transformation) was the *result* of two independent blind forces, a simple couple of natural forces allowing biological complexity in the long run. It cannot be over-emphasised that the successful introduction of stochastic concepts in biology would profoundly modify human thinking in all its aspects, and would contribute to prepare the minds to the strange concepts of quantum mechanics in the 20th century.

So now biologists had a coherent phylogenic model offering structural stability even within the framework of evolution, through the strictly vertical and time-oriented flow of information, which is simply represented by a genealogical tree (with only one possible way between any two *nodes*, rather than the many possible ways that may exist in a reticular network; the link between two adjacent nodes is properly named a *branch*). In this phylogenetic tree, evolution happens by speciation (divergence of populations and emergence of species) or extinction (by accidents of any kinds, or by natural selection). Many biologists helped in the development of this model, among them ERNST HÆCKEL, who was the first, in 1866, to propose a phyletic tree of all living organisms.

Inevitably, the next logical step in the direction of mechanistic reductionism would be to discover how exactly morphological characters could be modified within an organism, and how exactly these characters could be transmitted from one generation to another. This happened in a way demonstrating that scientific progress is not always (indeed rarely...) ordered and linear. Firstly, from Johann Gregor MENDEL's horticultural *hybridisation* experiments on *Pisum sativum*, remarkably exposed in 1866 ("Versuche über Pflanzenhybriden"), and apex, since the works of the botanist Joseph Gotlieb KÆLREUTER in the second half of the 18th century, of more than a century of hybridisation experiments; this was the work on which what would become to be known as the first laws of genetics would be founded... in 1899-1901 ! Secondly, from the pioneering cytological researches of OSCAR HERTWIG, EDWARD STRASBURGER, WALTHER FLEMMING, AUGUST WEISMANN, EDOUARD VAN BENEDEN, WILHELM WALDEYER, HUGO DE VRIES, among others, that contributed to show that *chromosomes* were the physical support of heredity.

At long last, the material support of transmission of hereditary characters, which had eluded researchers in life sciences since 2 millenaries and a half, could be traced to a particular kind of material inside the nucleus of some specialised cells, and the basic laws of heredity were known ! In the dawn of the 20th century, the new science of genetics was promisingly shining, and many great researchers helped to establish it as both an experimental and mathematical science: WILLIAM BATESON, LUCIEN CUÉNOT, WALTER S. SUTTON, THEODOR BOVERI, DARBISHIRE, F. A. JANSSENS, WILHELM JOHANNSEN, GODFREY HARDY, WILHELM WEINBERG, and others, and particularly THOMAS HUNT MORGAN. The morphological characters that helped to define taxonomic units were linked to theoretical units, named genes, that for the moment could not be seen as such, but of which one could predict with satisfactory precision the hereditary behaviour. Through the works of RONALD A. FISHER, SEWAL WRIGHT and J. B. S. HALDANE, these laws developed in a sophisticated mathematical framework, *populations genetics*.

### **The hard parts of phylogenetics**

The next great leap would happen from the precise molecular understanding of the heredity carriers within chromosomes, i.e. the *nucleic acids* (the DNA) composing genes, and of the biochemical machinery about it. This began 30 years ago, and now we know that the ultimate hereditary "atomon"-characters (which DEMOCRITUS was already thinking about) composing genes are *nucleotidic bases* which can appear as one of four possible states, and chemically bound together in a linear primary structure, thus forming genetic *sequences* which are in theory easily amenable to *homologous* comparisons (homologous sequences can be aligned in a matrix framework, with the characters forming the columns and the sequences forming the rows, on the basis that they are comparable enough objects that share a not too different common ancestor).

As a consequence, phylogenists now dispose of many different kinds of characters to analyse, from morphological to nucleotidic. After populations genetics, transformist

phylogenetics has also become a mathematically advanced domain of research, frequently difficult (and young enough, and sometimes sufficiently obscure, to be still liable to often acrimonious debates). These two quantitatively-inclined sciences basically treat of the same kind of problem, but with two fundamentally different visions, one being based on the individuals (the organisms) composing populations, the other on the natural groups (the taxa) that these populations form. As a result, the time scales are very different, as are the notions of the environment, of *anagenesis* (evolutionary line during which a character or a taxon is progressively modified, Gr. ana = up) and of phylogenesis, and because of this there is still a large practical and conceptual gap between populations genetics and taxonomic phylogenetics.

Even within taxonomic phylogenetics there can be large gaps between different “schools”, because of differences in the nature of the analysed characters on the one hand, in the mathematical tools of analysis on the other hand. Classically, phylogenesis can be inferred through the study of extraordinary (rare) events, such as macro-mutations, e.g. rearrangements of the genome. Generally, important morphogenetic changes can occur from *macro-mutations* (particularly through embryo- or epi-genesis, or embryological differentiation), hence the characters which are subject to such mutations often are *phenotypically* (Gr. phainein = to appear) expressed, and thus can sometimes be correlated to paleontological data, which is a powerful advantage for establishing lines of evolution.

However, the time flow as defined by these macro-mutations is strongly irregular in its discontinuity, and moreover phenotypic characters are strongly liable to *homoplasy* (Gr. homos = similar, and plasis, from plassein = to model, to shape), i.e. independent (relatively to lineages, not to causes) occurrences of similar characters, through *convergence (parallelism)* or *reversion*. As classical examples of homoplasy, one can mention active flight, that has developed at least twice in vertebrates, once in birds and once in bats, with the latter, despite what can be inferred from the Old Testament, being not birds, but mammals: though the *forelimbs* of birds and bats are Tetrapoda-specific homologous organs (expressed by a nexus of genes homologous to all tetrapods), their wings simply serve, through convergence within the Tetrapoda taxon, an *analogous* function. As an other example, one can also mention cetaceans and sirenians, which are mammals having – independently – returned to a water habitus, and which are not, despite superficial anatomical resemblance, some special kinds of fish. Another classical example is the eye: as a *photoreceptor* it is a Metazoa-specific homologous organ expressed by a nexus of genes homologous to all metazoans, but as a true eye, i.e. an organ able to produce a *complex* image, it has developed many times in evolution within the Metazoa taxon. From these examples, it is clear that homology is an eminently *relative* concept, which depends on the taxonomic level on refers to.

Simply said, two doubles do not necessarily share family ties as twins: their similarity may be an expression of homoplasy rather than homology, but in the world of anatomical phylogenesis this is not always easy to recognise. Consequently, it can

be very difficult to deduce phylogeny from phenotypes, unless there is palaeontological data to help; indeed, phenotypes can be quite misleading on the matter of phylogeny, and much effort is asked from evolutionary morphologists in terms of attention in their observations, and subtlety in their deductions.

The other way through which phylogenesis can be inferred is from the quantitative analysis of more frequent events (but not *too* frequent, so as to avoid quick saturation and loss of information). Now, within replicating organisms, nucleotidic (peptidic) chainmolecules do accumulate, more or less steadily, *point mutations* in their genes. These events (e.g. in proteins or, more precisely, in protein-coding genes), are micro-mutations, which allow, together with macro-mutations, evolution to happen, yet are more liable to numerical analysis than the latter. The source of this kind of data is: laboratory sequencing of these nucleotidic or proteinic sequences, which can now in a good part be automated, followed by homologous *alignments* of these sequences. It is important to note that, contrary to phenotypic data, molecular sequences can easily be stored in computerised data-bases for retrieval and further analysis. Despite the general absence of fossil genes (at least for the time being), micro-mutating molecular sequences have proved to be excellent sources of numerical information for phylogenetics, and their usefulness as a generally reliable and consistent source of information has been recognised for 30 years.

Moreover, during the last ten years, the improvements in biological molecular techniques have enabled sequences to be obtained more and more rapidly. This has resulted in a considerably increasing amount of available sequences. In parallel, micro-computers have become more and more powerful and easy to use. The recent explosion in the availability of molecular sequence data, and the simultaneously increasing accessibility of powerful processing tools, have led to a growing interest and use of computer models to compare homologous sequences for the purposes of inferring molecular evolution and phylogeny, i.e. not only the taxonomic evolution of natural groups of organisms (*orthologous* phylogeny, through speciation), but also the molecular evolution of genomes or families of genes (*paralogous* phylogeny, through gene duplication).

Consequently, more and more efficient quantitative analysis tools for data from molecular chains have been being elaborated, and still are, because the main problem of phylogenetics is not an easy one: how does one reliably reconstruct a phylogeny while having characters information for only the *terminal* nodes of a tree, i.e. the current sequences (the “*leaves*”) ?

Presently, dozens of phylogenetic programs or libraries of programs exist, which address this problem, some of the better known including Clustal, PAUP, and the Phylip library. So let us briefly see what are the main numerical methods generally used that allow the drawing of a phyletic *dendrogram* (Gr. dendron = tree, gramma = writing).

The first method is based on the principle of parsimony as applied to dissimilarities between contiguous nodes in the tree: the objective is to minimise the total sum of these dissimilarities. This is the method of choice of the *cladistic* (Gr. klados =

branch) school which, making a clear distinction between *ancestral* and *derived* states of character (relative, not absolute, notions, called *plesiomorphies* and *apomorphies* - Gr. plêsios = neighbour, apo = far), bases its *cladograms* only on homologous apomorphies (or *synapomorphies* - Gr. syn = with), with minimisation of homoplasy.

An interesting by-product, albeit an ambiguous and double-edged one, of these CLADISTIC MAXIMUM PARSIMONY (CMP) methods, is that they propose the reconstruction of the character states of hypothetical ancestors. The main drawbacks of these methods are: their *slowness* (this is true not only of the exhaustive search procedure, but also of the “branch and bound” algorithms which guarantee finding all the MP trees without proceeding with an exhaustive search of all possible trees); their *non consistency*, i.e. there is not necessarily convergence of outputs when the data on a given problem grows; and the necessity of *consensus* rules, because the proposed most parsimonious scenarios can be quite numerous, necessitating *consensus* rules. Moreover, because of their slowness, the CMP computer programs generally propose (often unpublished) *heuristic* algorithms (Gr. heuriskein = to find), to help the user in his search of the most parsimonious trees; these are, basically, ingenious automatisations of clever guessing, and despite their elegant and optimistic name one must be aware of the strong possibility that they can completely mislead the user to a *local* minimum (only part of the MP trees, a part that may be non representative of the whole, or even non-MP trees), rather than to the global minimum of all truly MP trees. Thus, to avoid too flagrant artefacts in the combinatorial search, much branch-swapping must be carried out during the stepwise addition performed by these algorithms, slowing down the process without yet completely guaranteeing the result. This is an important problem of which too many users are not really aware.

A CHARACTERS COMPATIBILITY method proposes a tree that allows, for the biggest possible number of characters (a “clique”), such a scenario that their state modifications could have happened without any case of homoplasy. These characters are said to be mutually compatible. This kind of approach can may be considered as a particular sort of CMP methods.

A third and quite diverse group of phylogenetic methods is formed by NUMERICAL TAXONOMIC PHENETICS (NTP) methods. As implied by their name, these are classification methods based on *global* similarities of compared units, more precisely, on the *matrix of dissimilarities* between all possible pairs of comparable units (sequences from the leaves, e.g. from terminal taxa in the case of an orthologous tree). There are many available NTP methods, some of them being simple clustering tools (e.g. UPGMA), others, such as NJ, being more sophisticated. NTP methods are not necessarily and *stricto sensu* phylogenetic methods, albeit phylogenesis can be inferred by the *phenograms* thus produced. The absence of phyletic *a priori* hypotheses can even be considered as an advantage under some conditions. Another, more clear, advantage of these methods, is their intrinsic *rapidity* which, particularly when a data set is large, can be very useful, at least in a first exploratory approach of the available data..

Yet, with classical methods of numerical phenetics, a common problem is that strongly differing rates of molecular divergence within the different lineages have the artificial result of joining together taxa resulting from slowly-evolving lineages, and of pushing taxa from quickly-evolving lineages (e.g. those rich in *autapomorphies* - Gr. auto = self) out to a basal (out-group) position in the phenogram. Moreover, homoplastic characters are not distinguished from homologous characters, and consequently phenomena of convergence and reversion are not taken into account : accordingly, the absence of distinction between synapomorphies and symplesiomorphies results in evolutionary reversible lineages being treated as slowly-evolving, and thus being artificially grouped. So a net result of classical phenetical methods, of which one must be aware, is to group together taxa sharing plesiomorphies (primitive characters) on one hand, and to group together convergent taxa (e.g. sharing common adaptive responses) on the other hand, rather than to group them according to common ancestry. Again, as with CMP methods, NTP users are not always clearly aware of this important problem.

Finally, there is the group of *probabilistic* methods, of which the subgroup of MAXIMUM LIKELIHOOD ESTIMATION (ML) methods is the most developed. This is clearly the most promising, but also the most complex, group of phylogenetic methods. Their most fundamental quality is that they require prior definition of an *explicit* model of characters evolution, thus allowing rigorous statistical appraisal of the computed results. On the other hand, these are basically very *slow* methods, even if the proposed model is relatively simple. Thus, despite their essential rigor, their slowness and conceptual complexity do not make them popular. In fact, despite their inherent quality, an important quantum leap in computer speed should happen for MLE methods to have an important impact on phylogenetic reconstruction, because available molecular sequences data are, for the present moment, growing at even a more rapid rate than the speed of computers does! For all these reasons, probabilistic methods surely are the methods of choice for the future, but when generalized parallel computing shall come of age.

In summary, phylogenetic numerical methods are conceptually and practically very powerful and useful. But they may consume significant computer memory and/or processor time, or may create artefacts in their reconstitution, or are not easily able to account for specific biochemical and natural selection phenomena. Nevertheless, there is continuous development in this sphere of research, and more than ever the synergic cooperation of naturalists, molecular biologists and quantitativists appears to be a good path towards one of the principal aims of science: understanding the processes and stories of life, understanding evolution.

## RÉSUMÉ

Les racines historiques de la phylogénétique sont brièvement présentées, de l'Antiquité au début du 20ème siècle. Puis le corps des concepts et méthodes techniques de cette science, qui est profondément enracinée et en pleine croissance, est succinctement décrit.

## SHORT BIBLIOGRAPHY

- BITTAR, Gabriel. 1995. "Phylogénie: histoire, concepts, principes et présentation de nouvelles méthodes d'analyse numérique de l'évolution moléculaire". Thèse de doctorat n°2789, Faculté des Sciences de l'Université de Genève, 480 pp.
- Collectif, sous la dir. de J.-P. ICIKOVICS, Décembre 1991. "Darwin ou Lamarck. La Querelle de l'Evolution". *Les Cahiers de Science et Vie* n°6, Paris, 96 pp.
- Collectif, sous la dir. de J.-P. ICIKOVICS, Juin 1993. "Mendel. Comment naquit la génétique". *Les Cahiers de Science et Vie* n°15, Paris, 96 pp.
- DARLU, Pierre, & TASSY, Pascal, 1993. "La reconstruction phylogénétique. Concepts et méthodes". Masson, Paris, 245 pp.
- DAUDIN, Henri, 1926. "De Linné à Lamarck. Méthodes de la classification et idée de série en botanique et en zoologie (1740-1790)". *Presses Universitaires de France, Paris; réimpression Editions des Archives Contemporaines, Montreux, Paris, 1983, 264 pp.*
- DESMOND, Adrian, & MOORE, James, 1991. "Darwin". *Michael Joseph, London, 808 pp.*
- DOOLITTLE, R. F., editor, 1990. "Molecular evolution: computer analysis of protein and nucleic acid sequences". *Methods in Enzymology* vol. 183, Academic Press, San Diego, 736 pp.
- DROUIN, J.-M., & LENAY, Ch., éditeurs, 1990. "Théories de l'évolution. Une anthologie". *Presses Pocket, Paris, 214 pp.*
- DE WIT, Hendrik C. D., 1982-9. "Histoire du développement de la biologie". Trad. du hollandais par A. Baudière, *Presses polytechniques et universitaires romandes, 3 vol., 1992-4, 404, 460 et 635 pp.*
- GREEN, Peter, 1990. "Alexander to Actium. The historical evolution of the Hellenistic Age". *University of California Press, Berkeley, Los Angeles, 970 pp.*
- HILLIS, D. M., & MORITZ, CR., & MABLE, B.K., editors, 1990. "Molecular Systematics". 2nd Edition. *Sinauer Associates, Sunderland, 588 pp.*
- LLOYD, Geoffrey E. R., 1970. "Early Greek Science: Thales to Aristotle". *Norton & Co, New York, London, 156 pp.*
- LLOYD, Geoffrey E. R., 1973. "Greek Science after Aristotle". *Norton & Co, New York, London, 189 pp.*
- LLOYD, Geoffrey E. R., 1979. "Magic, reason and experience. Studies in the origin and development of Greek science". *Cambridge University Press, Cambridge, 335 pp.*
- LLOYD, Geoffrey E. R. 1983. "Science, folklore and ideology. Studies in the life sciences in Ancient Greece". *Cambridge University Press, Cambridge, 260 pp.*
- MIYAMOTO, M. M., & CRACRAFT, J., editors, 1991. "Phylogenetic analysis of DNA sequences". *Oxford University Press, Oxford, New York, 359 pp.*
- PACAUD, Adolphe, 1995. "Du feu". *Hachette, Paris, 108 pp.*
- VOILQUIN, J. 1964. "Les penseurs grecs avant Socrate. De Thalès de Milet à Prodicos". *Garnier-Flammarion, Paris, 247 pp.*