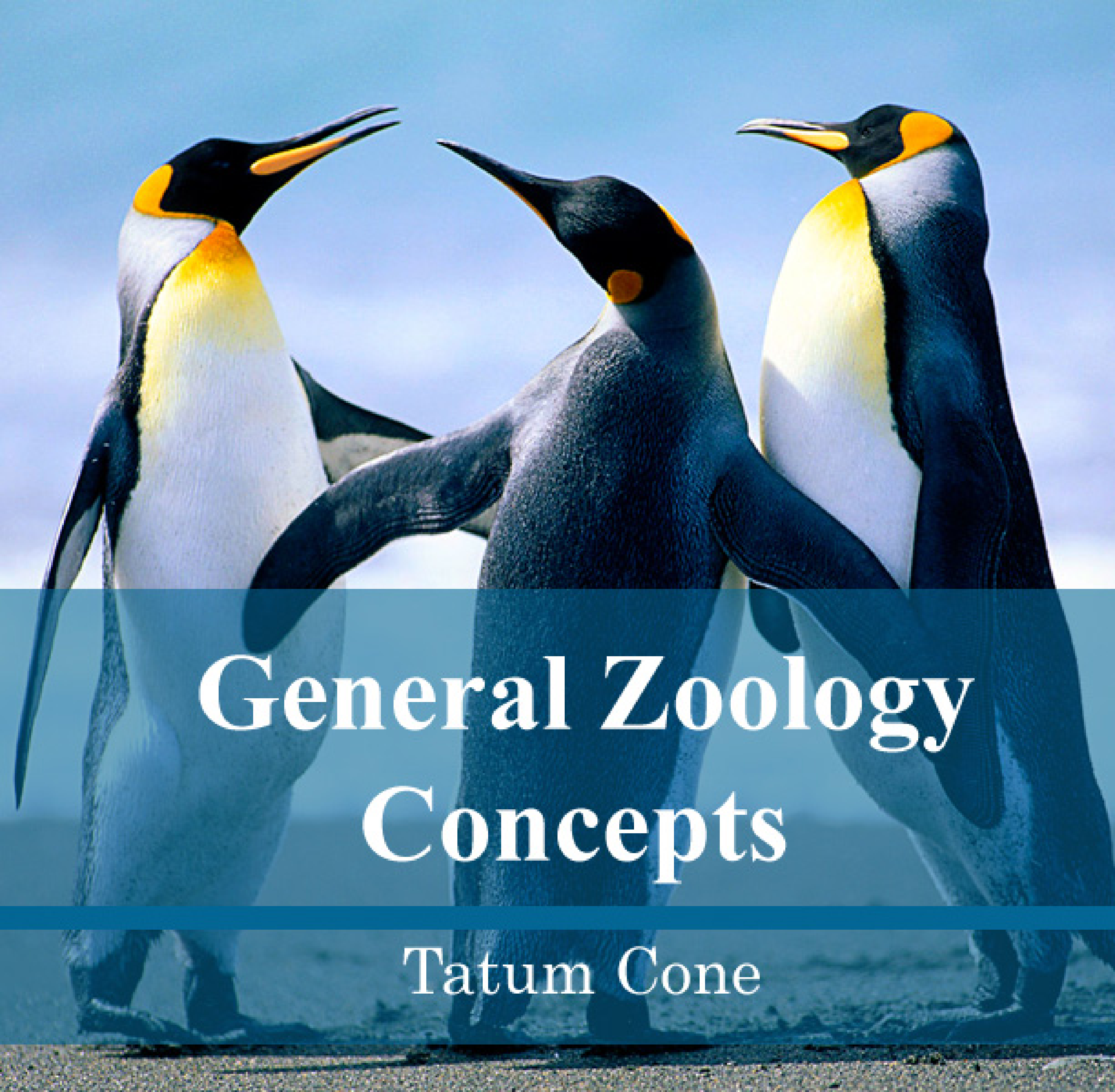


General Zoology Concepts

Tatum Cone

First Edition, 2012

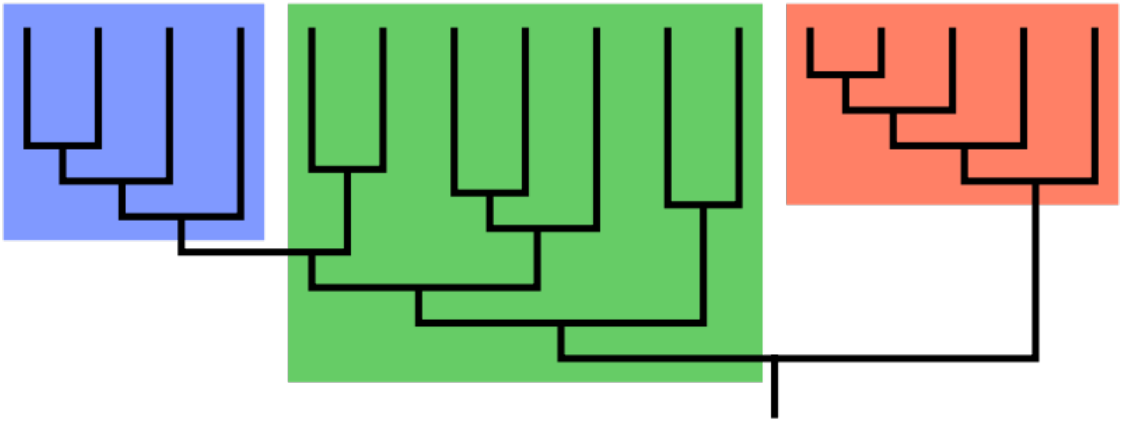
ISBN 978-81-323-3148-3

Published by:
Research World
4735/22 Prakashdeep Bldg,
Ansari Road, Darya Ganj,
Delhi - 110002
Email: info@wtbooks.com

Table of Contents

Chapter- 1

Clade

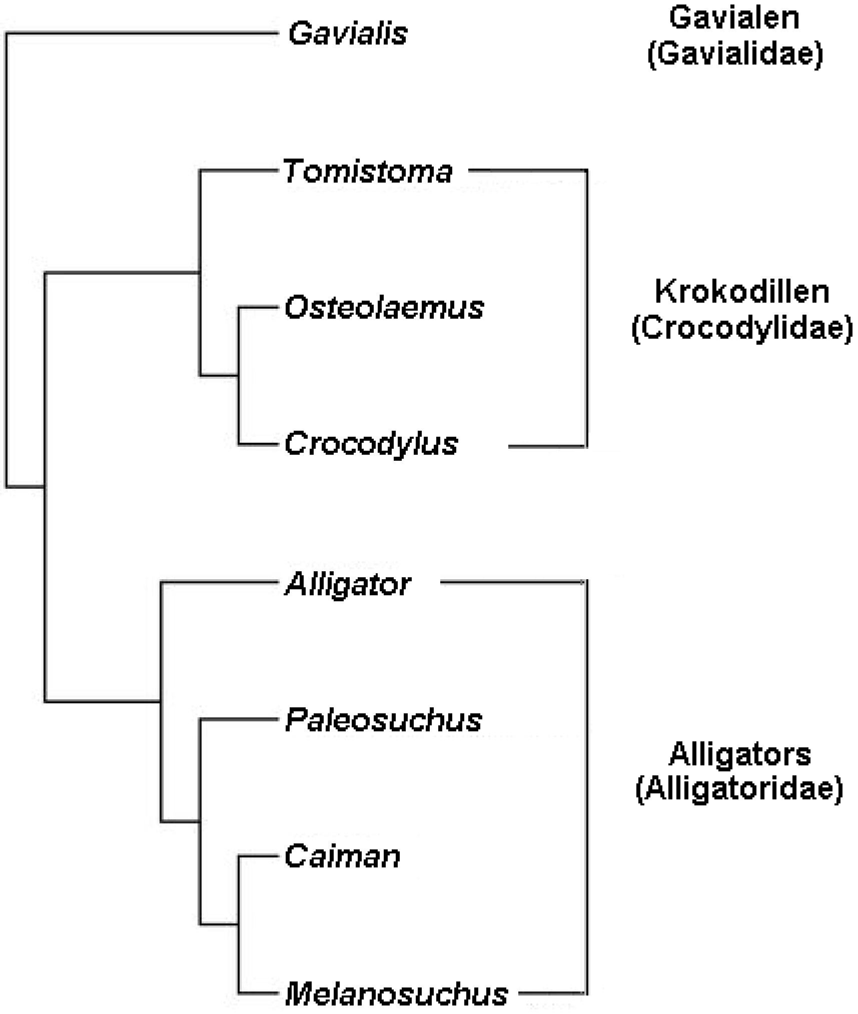


Cladogram (family tree) of a biological group. The red and blue boxes represent *clades* (i.e., complete branches). The green box is not a clade, but rather represents an *evolutionary grade*, an incomplete group, because the blue clade descends from it, but is excluded.

A **clade** is a group consisting of an organism and all its descendants. In the terms of biological systematics, a clade is a single "branch" on the "tree of life". The idea that such a "natural group" of organisms should be grouped together and given a taxonomic name is central to biological classification. In cladistics (which takes its name from the term), clades are the only acceptable units.

The term was coined in 1958 by English biologist Julian Huxley.

Definitions



A cladogram of crocodiles, a visual representation of their relationship

Clade and ancestor

A clade is termed monophyletic, meaning it contains one ancestor which can be an organism, population, or species and all its descendants. The term clade refers to the grouping of the ancestor and its living and/or deceased descendants together. The ancestor can be a theoretical or actual species.

Clade definition

Three methods of defining clades are featured in phylogenetic nomenclature: node-, stem-, and apomorphy-based:

- In node-based definition, clade name A refers to the *least inclusive* clade containing taxa (or specimens) X, Y, etc., and their common ancestor. The ancestor is the branch point, or *node*.
- In stem-based definition, A refers to the *most inclusive* clade containing X, Y, etc., and their common ancestor, down to where Z branches off below A. Taxa are included between the node of A and down to (but not including) the branching point to Z; that is, the *stem* of A.
- In apomorphy-based definition, A refers to the clade identified by an apomorphy (a trait) found in X, Y, etc., and their common ancestor.

In Linnaean taxonomy, clades are defined by a set of traits (apomorphies) unique to the group. This system is basically similar to the apomorphy-based clades of phylogenetic nomenclature. The difference is one of weight: While phylogenetic nomenclature bases the group on an ancestor with a certain trait, Linnaean taxonomy uses the traits themselves to define the group.

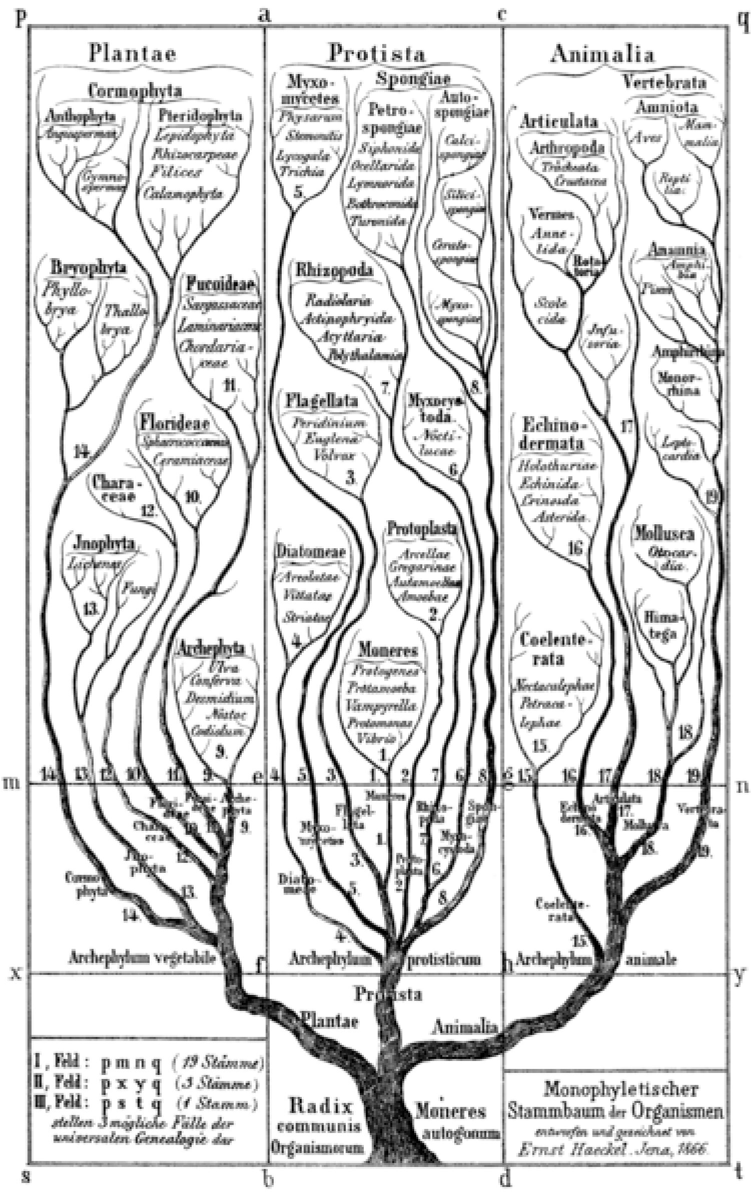
Clades as constructs

In cladistics, the clade is a hypothetical construct based on experimental data. Clades are found using multiple (sometimes hundreds) of traits from a number of species (or specimens) and analysing them statistically to find the most likely phylogenetic tree for the group. Although similar in some ways to a biological classification of species, the method is statistical and more open to scrutiny than traditional methods. Although taxonomists use clades as a tool in classification where feasible, the taxonomic "tree of life" is not the same as the cladistic. The traditional genus, family, etc. names are not necessarily clades; though they will often be.

Clade names

In Linnaean systematics, the various groups are ordered into a series of taxonomic ranks (the familiar order, family etc.). These ranks will by convention dictate the ending to names for some groups. Clades do not by their nature fit this scheme, and no such restriction exists as to their names in cladistics. There is however a convention for naming more or less inclusive groups, which are given prefixes like *crown-* or *pan*.

Taxonomy and systematics



Early phylogenetic tree by Haeckel, 1866

The idea of a "clade" did not exist in pre-Darwinian Linnaean taxonomy, which was based only on morphological similarities between organisms – although it happens that many of the better known animal groups in Linnaeus' original Systema Naturae (notably among the vertebrate groups) do represent clades. With the publication of Darwin's theory of evolution in 1859, taxonomy gained a theoretical basis, and the idea was born that groups used in a system of classification should represent branches on the evolutionary tree of life. In the century and a half since then, taxonomists have worked to make the taxonomic system reflect evolution. However, partly because the Tree of Life branches rather unevenly, the hierarchy of the Linnaean system does not always lend itself well to representing clades. The result is that when it comes to naming, cladistics and Linnaean taxonomy are not always compatible. In particular, higher level taxa in Linnaean taxonomy often represent evolutionary grades rather than clades, resulting in groups made up of clades where one or two sub-branches have been excluded. Typical examples include bony fishes, which include the ancestor of tetrapods, and reptiles, ancestral to both birds and mammals.

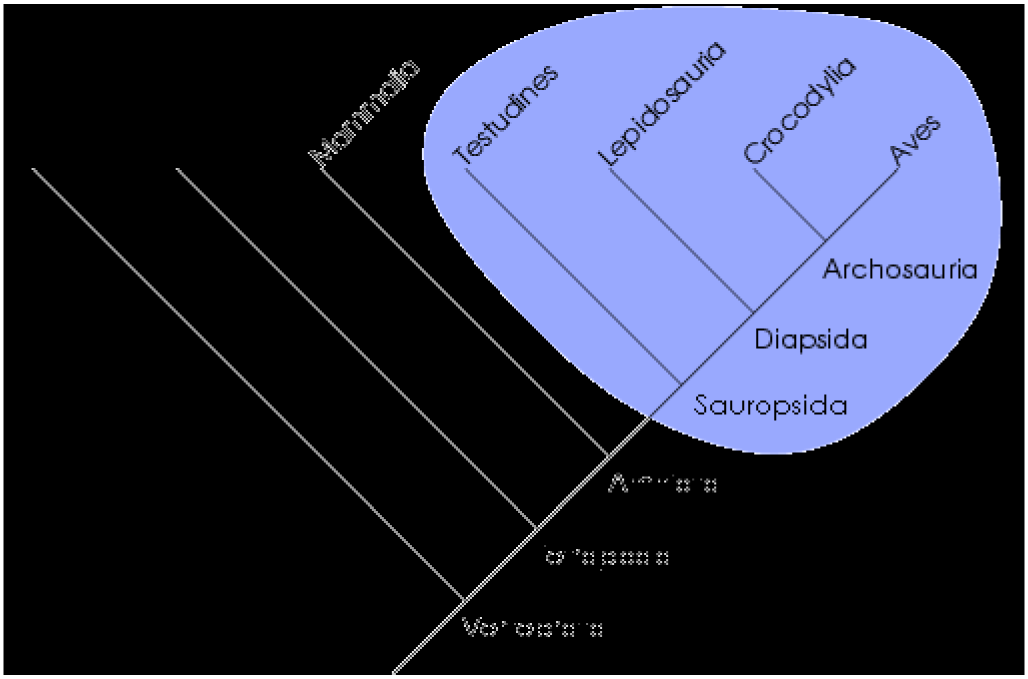
In phylogenetic nomenclature, clades can be nested at any level, and do not have to be slotted into a small number of ranks in an overall hierarchy. In contrast, the Linnaean units of "order", "class" etc. must be used when naming a new taxon. As there are only seven formal levels to the Linnaean system (species being the lowest), only a finite number of sub- and super-units can be created. In order to be able to use the full complexity of taxonomic trees (cladograms) in an area with which they are very familiar, some researchers have opted to dispense with ranks all together, instead using clade names without Linnaean ranks. The reason for preferring one system over the other is partly one of application: cladistic trees give details, suitable for specialists; the Linnaean system gives a well ordered overview, at the expense of details of the phylogenetic tree.

In a few instances, the Linnaean system has actually impeded our understanding of the phylogeny and broad evolutionary patterns. The best known example is the interpretation of the strange fossils of the Burgess Shale and the subsequent idea of a "Cambrian Explosion" With the application of cladistics, and the rejection of any significance of the concept of phyla, the confusion of the late 20th century over the Burgess animals has been resolved. It appears there never was an "explosion" of major bauplans with subsequent extinctions. The seemingly weird critters themselves have been found to be representatives of a group, the Lobopodia, that includes arthropods, water bears and velvet worms.

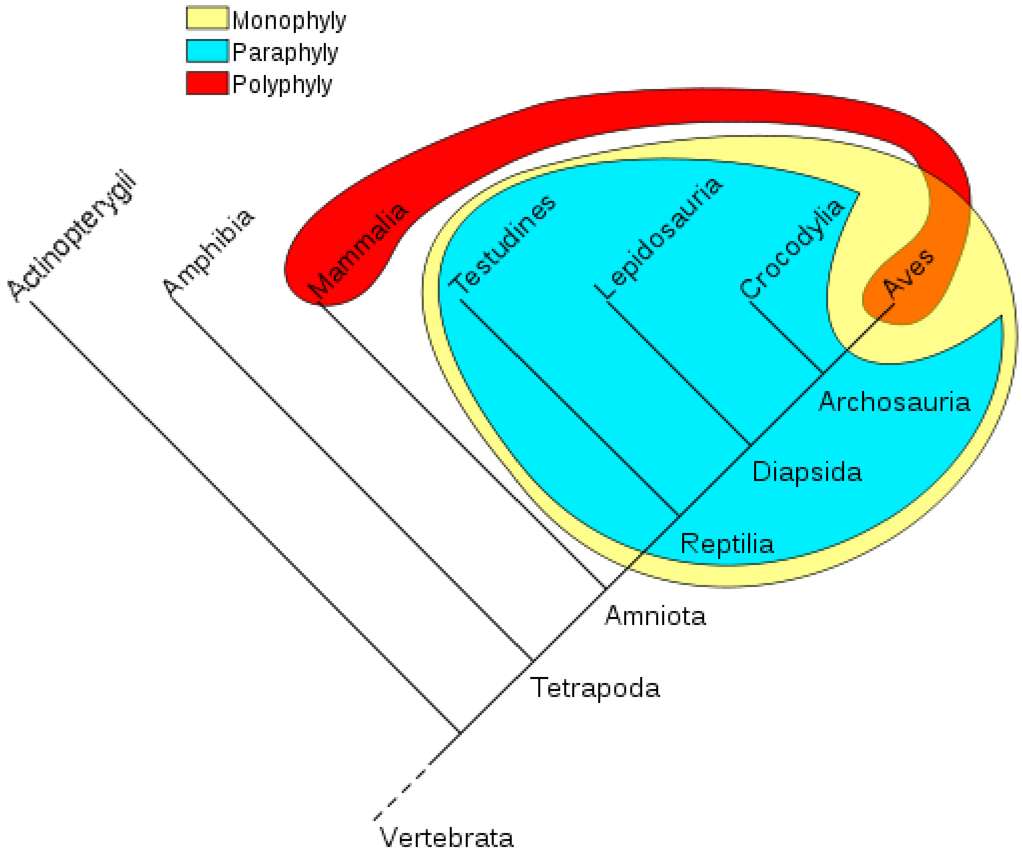
In most instances the two systems are not at odds, however. The cladistic statement, that the clade Lobopodia contains (among others) the Arthropoda, Tardigrada and Onychophora, is factually identical to the Linnaean evolutionary statement that the group Lobopodia is ancestral to the phyla Arthropoda, Tardigrada and Onychophora. The difference is one of semantics rather than phylogeny.

Chapter- 2

Monophyly



A monophyly



Partial Evolutionary Tree of the Vertebrates

Comparison of phylogenetic groups, showing a **monophyly** (all descendants of the first reptiles), a **paraphyly** (descendants of reptiles, minus birds), and a **polyphyly** (warm-blooded animals: mammals and birds)

In common cladistic usage, a **monophyletic** group is a taxon (group of organisms) which forms a clade, meaning that it contains *all* the descendants of the possibly hypothetical closest common ancestor of the members of the group. The term is synonymous with the uncommon term holophyly. Monophyletic groups are typically characterized by shared derived characteristics (synapomorphies).

Monophyly is contrasted with the terms paraphyly and polyphyly, which are most easily understood from the second diagram. In current usage, a **paraphyletic** group consists of all of the descendants of a possibly hypothetical closest common ancestor *minus* one or more monophyletic groups (most usually one). A paraphyletic group is thus 'nearly' monophyletic (consistent with the meaning of the prefix 'para', namely 'near' or 'alongside'.) A **polyphyletic** group is any group other than a monophyletic group or a paraphyletic group, which like a paraphyletic group contains only some of the descendants of their closest common ancestor, but unlike a paraphyletic group is not characterized by the missing descendants forming one (or more) monophyletic groups.

These definitions have taken some time to be accepted. When the cladistic school of thought became mainstream in the 1960s, several alternative definitions were in use. Indeed, taxonomists sometimes used terms without defining them, leading to confusion in the early literature, a confusion which persists.

Definitions

On the broadest scale, definitions fall into two groups.

- The widest, and arguably the semantically correct meaning of the word, is any two or more groups sharing a common ancestor. This very broad definition strips the term of scientific utility. Therefore, scientists today restrict the term to holophyletic groups only – that is, groups consisting of all the descendants of one (usually hypothetical) common ancestor. However, when considering taxonomic groups such as genera and species, the most appropriate nature of their common ancestor is unclear. Assuming that it would be one individual or mating pair is unrealistic for sexually reproducing species, which are by definition interbreeding populations.

- However, using a broader definition, such as a species and all its descendants, does not really work to define a genus. A satisfactory and comprehensive cladistic definition of a species or genus is in fact impossible, and reflects the impossibility of seamlessly impressing a gradualistic model of continual change over the 'quantum' Linnean model, where species have defined boundaries, and intermediaries between species cannot be accommodated.

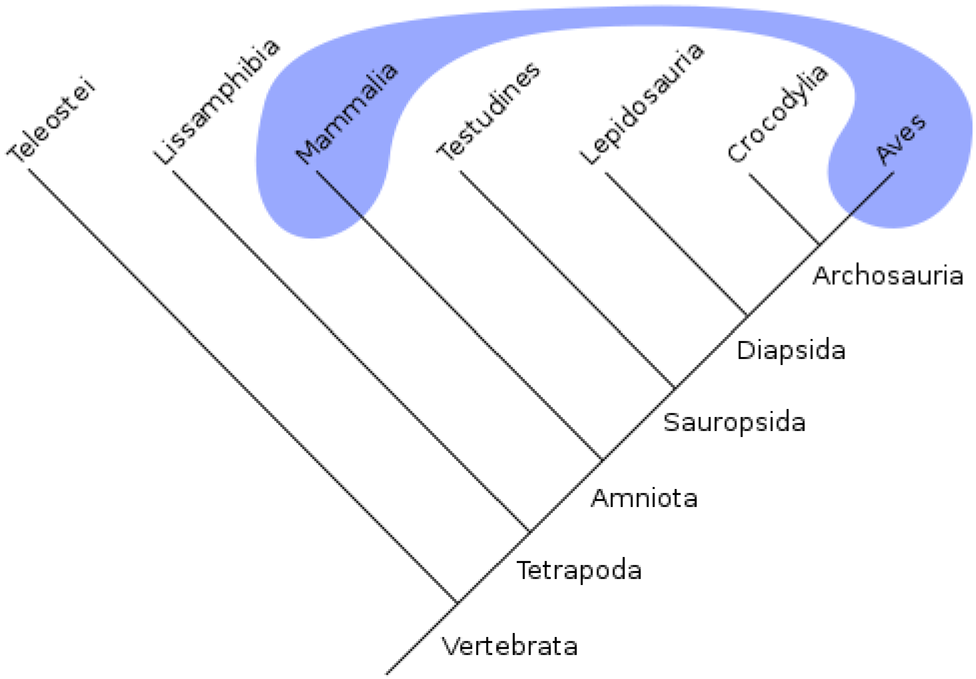
Controversy

This incompatibility with the Linnaean taxonomy model led to an initial rift, not entirely healed, between the cladistic and Linnean schools of thought. Extreme cladists challenged the validity of Linnean taxa such as the Reptilia. Because birds, although descended from reptiles, are not themselves considered to be reptiles, cladists demanded that the taxon Reptilia be dismantled: a request that taxonomists were unwilling to heed. This stand-off was eventually resolved to a degree by the construction of the term 'paraphyletic' to describe closely related groups which included most but not all of the descendants of a common ancestor.

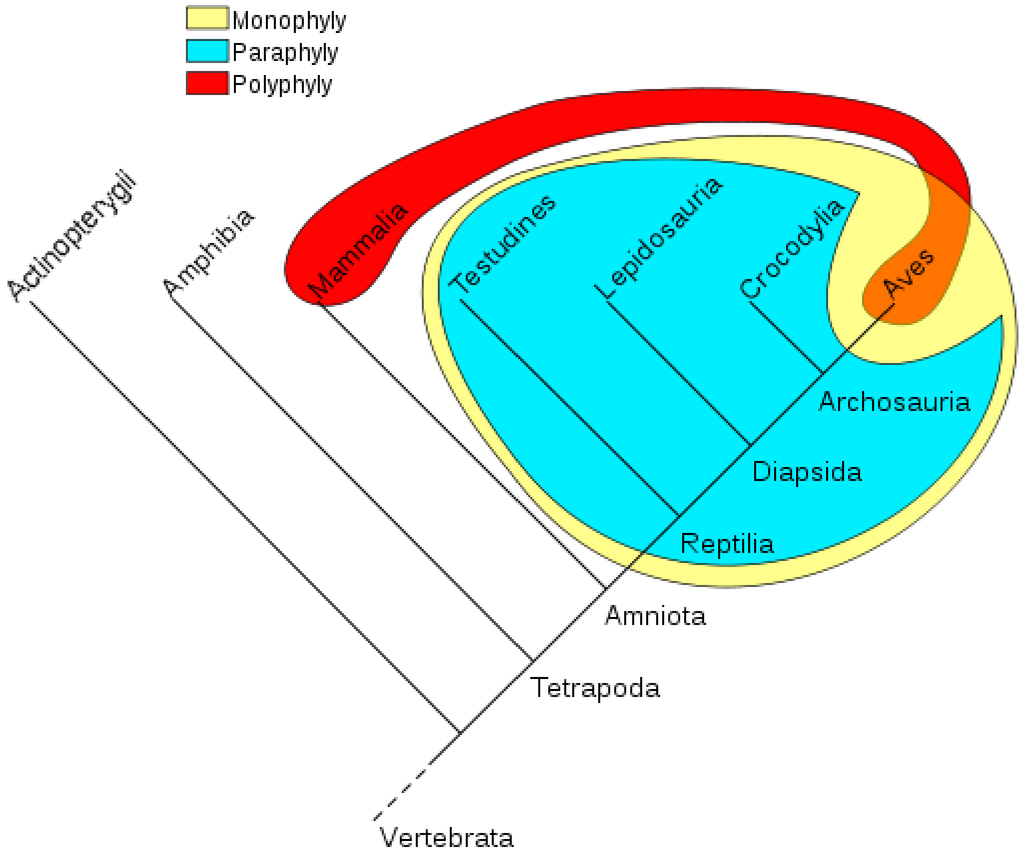
However, the coining of this term led to yet more confusion. Some scientists considered paraphyletic groups to be monophyletic (as they shared a common ancestor), where others insisted that monophyletic should continue to refer only to holophyletic groups. Another term, polyphyletic, fell outside of the definition of monophyly. A strict explanation of a paraphyletic group has not been published, but the consensus appears to be that paraphyletic groups consist of a monophyletic group, minus one smaller constituent clade – for instance "reptiles minus birds". Polyphyletic groups can be thought of as a number of unrelated clades, for instance "warm blooded animals" = "birds plus mammals". Non-holophyletic groups are of little use for analysis of evolutionary processes, hence the calls for their "unnaming" - even though they are useful to scientists who are less concerned with the evolutionary past of groups. Naming is also a problem for monophyletic groups: because the number of ancestors from which to root monophyletic groups is almost infinite, giving each clade a unique name is impossible - as illustrated by the failed attempts to instigate a system called the Phylocode. Names obfuscate the really interesting part, which is the branching order, and are therefore of little utility to the cladist - at odds with the taxonomist, who since the time of Linnaeus has been naming species. Intermediate, and particularly fossil, taxa can be considered to fall 'just outside' a widely accepted taxon. For instance, *Archaeopteryx* appears more reptilian than bird – it has teeth and a number of other reptilian characteristics. But it also has feathers, which have traditionally been considered as an avian trait. It lacks a number of other traits shared by all birds, so cannot fall within the bird clade. To reflect this phylogenetic proximity, it is termed a 'stem group bird' - i.e. it lies on a branch close to the lineage that led to true birds, as recognised by a taxonomist. This concept closes the gap between taxonomy and cladistics at a broader scale, but is difficult to apply at a species-level resolution.

Chapter- 3

Polyphyly



The group of "warm-blooded animals" is polyphyletic.



Partial Evolutionary Tree of the Vertebrates

Comparison: monophyly = yellow, polyphyly = red, paraphyly = aqua

A **polyphyletic** (Greek for "of many races") group is one whose members' last common ancestor is not a member of the group.

For example, the group consisting of warm-blooded animals is polyphyletic, because it contains both mammals and birds, but the most recent common ancestor of mammals and birds was cold-blooded. Warm-bloodedness evolved separately in the ancestors of mammals and the ancestors of birds, so it is not a true phylogenetic grouping.

Scientific classification aims to group species together such that every group is descended from a single common ancestor and therefore it is frequently a goal to eliminate groups that are found to be polyphyletic. This is often the stimulus for major revisions of the classification schemes. A polyphyletic group can be "fixed" either by excluding clades or by adding the common ancestor.

Opinions differ as to whether valid groups need to contain *all* the descendants of a common ancestor. Groups that do so are called monophyletic, and according to cladistics it should be the aim of classification to ensure that all groups have this property. However, many other taxonomists would argue that there is a valid place for groups that are paraphyletic, i.e. contains its most recent common ancestor but does not contain all the descendants of that ancestor.

Cladistics generally discourages polyphyletic groups

In most cladistics-based schools of taxonomy, the existence of polyphyletic groups (as well as paraphyletic groups) in a classification is discouraged. Monophyletic groups (that is, clades) are considered by these schools of thought to be the most important grouping of organisms, for the following reasons:

- Clades are simple to define: a typical clade definition is "All descendants of the nearest common ancestor of species X and Y". On the other hand, polyphyletic and paraphyletic groups are always defined in terms of clades, for example "reptiles are the Sauropsid clade, minus the Aves clade". Or "Warm-blooded animals are the Aves clade plus the Mammals clade". Because polyphyletic and paraphyletic groups are defined in terms of clades plus or minus other clades, they are considered less important than monophyletic (single, whole) clades.

- For a given evolutionary tree of, say, N nodes, there are exactly N clades (one per node). However, the number of paraphyletic groups and polyphyletic groups is exponentially larger than that, on the order of 2^N. Yet only a small fraction of the paraphyletic groups are given names or discussed.

- Polyphyletic groups often have their origin in traditional taxonomy, based on similar morphological characteristics. The original perception may have been that the group was entirely descended from a single ancestor. If such a group is later discovered (for instance, due to convergent evolution) to be polyphyletic, rather than monophyletic, then such a group loses its original significance.

Chapter- 4

Speciation

Speciation is the evolutionary process by which new biological species arise. The biologist Orator F. Cook seems to have been the first to coin the term 'speciation' for the splitting of lineages or 'cladogenesis,' as opposed to 'anagenesis' or 'phyletic evolution' occurring within lineages. Whether genetic drift is a minor or major contributor to speciation is the subject matter of much ongoing discussion.

There are four geographic modes of speciation in nature, based on the extent to which speciating populations are geographically isolated from one another: allopatric, peripatric, parapatric, and sympatric. Speciation may also be induced artificially, through animal husbandry or laboratory experiments. Observed examples of each kind of speciation are provided throughout.

Natural speciation

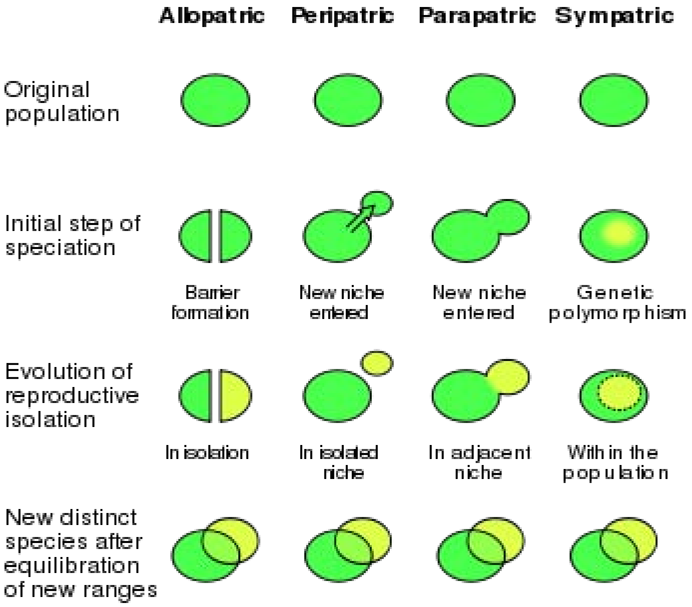
All forms of natural speciation have taken place over the course of evolution; however it still remains a subject of debate as to the relative importance of each mechanism in driving biodiversity.

The three-spined stickleback (*Gasterosteus aculeatus*)

One example of natural speciation is the diversity of the three-spined stickleback, a marine fish that, after the last ice age, has undergone speciation into new freshwater colonies in isolated lakes and streams. Over an estimated 10,000 generations, the sticklebacks show structural differences that are greater than those seen between different genera of fish including variations in fins, changes in the number or size of their bony plates, variable jaw structure, and color differences.

There is debate as to the rate at which speciation events occur over geologic time. While some evolutionary biologists claim that speciation events have remained relatively constant over time, some palaeontologists such as Niles Eldredge and Stephen Jay Gould have argued that species usually remain unchanged over long stretches of time, and that speciation occurs only over relatively brief intervals, a view known as *punctuated equilibrium*.

Allopatric



Comparison of allopatric, peripatric, parapatric and sympatric speciation.

During allopatric (from the ancient Greek *allos*, "other" + Greek *patrā*, "fatherland") speciation, a population splits into two geographically isolated populations (for example, by habitat fragmentation due to geographical change such as mountain building or social change such as emigration). The isolated populations then undergo genotypic and/or phenotypic divergence as: (a) they become subjected to dissimilar selective pressures; (b) they independently undergo genetic drift; (c) different mutations arise in the two populations. When the populations come back into contact, they have evolved such that they are reproductively isolated and are no longer capable of exchanging genes.

Observed instances

Island genetics, the tendency of small, isolated genetic pools to produce unusual traits, has been observed in many circumstances, including insular dwarfism and the radical changes among certain famous island chains, for example on Komodo. The Galápagos islands are particularly famous for their influence on Charles Darwin. During his five weeks there he heard that Galápagos tortoises could be identified by island, and noticed that Mockingbirds differed from one island to another, but it was only nine months later that he reflected that such facts could show that species were changeable. When he returned to England, his speculation on evolution deepened after experts informed him that these were separate species, not just varieties, and famously that other differing Galápagos birds were all species of finches. Though the finches were less important for Darwin, more recent research has shown the birds now known as Darwin's finches to be a classic case of adaptive evolutionary radiation.

Peripatric

In peripatric speciation, a subform of allopatric speciation, new species are formed in isolated, smaller peripheral populations that are prevented from exchanging genes with the main population. It is related to the concept of a founder effect, since small populations often undergo bottlenecks. Genetic drift is often proposed to play a significant role in peripatric speciation.

Observed instances

- Mayr bird fauna
- The Australian bird *Petroica multicolor*
- Reproductive isolation occurs in populations of *Drosophila* subject to population bottlenecking

The London Underground mosquito is a variant of the mosquito *Culex pipiens* that entered in the London Underground in the nineteenth century. Evidence for its speciation include genetic divergence, behavioral differences, and difficulty in mating.

Parapatric

In parapatric speciation, there is only partial separation of the zones of two diverging populations afforded by geography; individuals of each species may come in contact or cross habitats from time to time, but reduced fitness of the heterozygote leads to selection for behaviours or mechanisms that prevent their inter-breeding. Parapatric speciation is modelled on continuous variation within a 'single', connected habitat acting as a source of natural selection rather than the effects of isolation of habitats produced in peripatric and allopatric speciation.

Ecologists refer to parapatric and peripatric speciation in terms of ecological niches. A niche must be available in order for a new species to be successful.

Observed instances

- Ring species
    - The *Larus* gulls form a ring species around the North Pole.
    - The *Ensatina* salamanders, which form a ring round the Central Valley in California.
    - The Greenish Warbler (*Phylloscopus trochiloides*), around the Himalayas.
- the grass *Anthoxanthum* has been known to undergo parapatric speciation in such cases as mine contamination of an area.

Sympatric

Sympatric speciation refers to the formation of two or more descendant species from a single ancestral species all occupying the same geographic location.

In sympatric speciation, species diverge while inhabiting the same place. Often-cited examples of sympatric speciation are found in insects that become dependent on different host plants in the same area. However, the existence of sympatric speciation as a mechanism of speciation is still hotly contested. People have argued that the evidences of sympatric speciation are in fact examples of micro-allopatric, or heteropatric speciation. The most widely accepted example of sympatric speciation is that of the cichlids of Lake Nabugabo in East Africa, which is thought to be due to **sexual selection**.

Until recently, there has a been a dearth of strong evidence that supports this form of speciation, with a general feeling that interbreeding would soon eliminate any genetic differences that might appear. But there has been at least one recent study that suggests that sympatric speciation has occurred in Tennessee cave salamanders.

The three-spined sticklebacks, freshwater fishes, that have been studied by Dolph Schluter (who received his Ph.D. for his work on Darwin's finches with Peter J. Grant) and his current colleagues in British Columbia, were once thought to provide an intriguing example best explained by sympatric speciation. Schluter and colleagues found:

- Two different species of three-spined sticklebacks in each of five different lakes
    - a large benthic species with a large mouth that feeds on large prey in the littoral zone
    - a smaller limnetic species — with a smaller mouth — that feeds on the small plankton in open water
- DNA analysis indicates that each lake was colonized independently, presumably by a marine ancestor, after the last ice age
- DNA analysis also shows that the two species in each lake are more closely related to each other than they are to any of the species in the other lakes
- The two species in each lake are reproductively isolated; neither mates with the other.

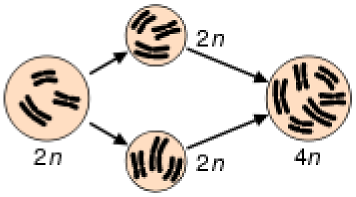
- However, aquarium tests showed:

    - the benthic species from one lake will spawn with the benthic species from the other lakes and
    - likewise the limnetic species from the different lakes will spawn with each other.
    - These benthic and limnetic species even display their mating preferences when presented with sticklebacks from Japanese lakes; that is, a Canadian benthic prefers a Japanese benthic over its close limnetic cousin from its own lake.
- Their conclusion: in each lake, what began as a single population faced such competition for limited resources that:
    - disruptive selection — competition favoring fishes at either extreme of body size and mouth size over those nearer the mean — coupled with:
    - assortative mating — each size preferred mates like it — favored a divergence into two subpopulations exploiting different food in different parts of the lake.
    - The fact that this pattern of speciation occurred the same way on three separate occasions suggests strongly that ecological factors in a sympatric population can cause speciation.

However, the DNA evidence cited above is from mitochondrial DNA (mtDNA), which can often move easily between closely related species ("introgression") when they hybridize. A more recent study, using genetic markers from the nuclear genome, shows that limnetic forms in different lakes are more closely related to each other (and to marine lineages) than to benthic forms in the same lake. The three-spine stickleback is now usually considered an example of "double invasion" (a form of allopatric speciation) in which repeated invasions of marine forms have subsequently differentiated into benthic and limnetic forms. The three-spine stickleback provides an example of how molecular biogeographic studies that rely solely on mtDNA can be misleading, and that consideration of the genealogical history of alleles from multiple unlinked markers (i.e. nuclear genes) is necessary to infer speciation histories.

Sympatric speciation driven by ecological factors may also account for the extraordinary diversity of crustaceans living in the depths of Siberia's Lake Baikal.

Speciation via polyploidization



Speciation via polyploidy: A diploid cell undergoes failed meiosis, producing diploid gametes, which self-fertilize to produce a tetraploid zygote.

Polyploidy is a mechanism often attributed to causing some speciation events in sympatry. Not all polyploids are reproductively isolated from their parental plants, so an increase in chromosome number may not result in the complete cessation of gene flow between the incipient polyploids and their parental diploids.

Polyploidy is observed in many species of both plants and animals, and results in rapid speciation since offspring of, for example, tetraploid x diploid matings result in triploid sterile progeny. It has been proposed that many of the existing plant and most animal species have undergone an event of polyploidization in their evolutionary history. However, reproduction is often by parthenogenesis since polyploid animals are often sterile. [needs further editing—not true in plants]. Rare instances of polyploid mammals are known, but most often result in prenatal death.

Hawthorn fly

One example of evolution at work is the case of the hawthorn fly, *Rhagoletis pomonella*, also known as the apple maggot fly, which appears to be undergoing sympatric speciation. Different populations of hawthorn fly feed on different fruits. A distinct population emerged in North America in the 19th century some time after apples, a non-native species, were introduced. This apple-feeding population normally feeds only on apples and not on the historically preferred fruit of hawthorns. The current hawthorn feeding population does not normally feed on apples. Some evidence, such as the fact that six out of thirteen allozyme loci are different, that hawthorn flies mature later in the season and take longer to mature than apple flies; and that there is little evidence of interbreeding (researchers have documented a 4-6% hybridization rate) suggests that sympatric speciation is occurring. The emergence of the new hawthorn fly is an example of evolution in progress.

Reinforcement (Wallace effect)

Reinforcement is the process by which natural selection increases reproductive isolation. It may occur after two populations of the same species are separated and then come back into contact. If their reproductive isolation was complete, then they will have already developed into two separate incompatible species. If their reproductive isolation is incomplete, then further mating between the populations will produce hybrids, which may or may not be fertile. If the hybrids are infertile, or fertile but less fit than their ancestors, then there will be no further reproductive isolation and speciation has essentially occurred (e.g., as in horses and donkeys.) The reasoning behind this is that if the parents of the hybrid offspring each have naturally selected traits for their own certain environments, the hybrid offspring will bear traits from both, therefore would not fit either ecological niche as well as either parent. The low fitness of the hybrids would cause selection to favor assortative mating, which would control hybridization. This is sometimes called the Wallace effect after the evolutionary biologist Alfred Russel Wallace who suggested in the late 19th century that it might be an important factor in speciation. If the hybrid offspring are more fit than their ancestors, then the populations will merge back into the same species within the area they are in contact.

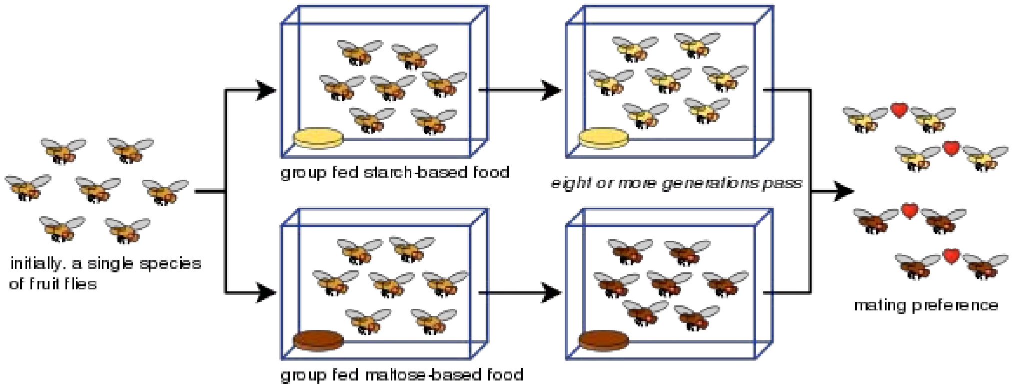
Reinforcement is required for both parapatric and sympatric speciation. Without reinforcement, the geographic area of contact between different forms of the same species, called their "hybrid zone," will not develop into a boundary between the different species. Hybrid zones are regions where diverged populations meet and interbreed. Hybrid offspring are very common in these regions, which are usually created by diverged species coming into secondary contact. Without reinforcement the two species would have uncontrollable inbreeding. Reinforcement may be induced in artificial selection experiments as described below.

Artificial speciation

New species have been created by domesticated animal husbandry, but the initial dates and methods of the initiation of such species are not clear. For example, domestic sheep were created by hybridisation, and no longer produce viable offspring with *Ovis orientalis*, one species from which they are descended. Domestic cattle, on the other hand, can be considered the same species as several varieties of wild ox, gaur, yak, etc., as they readily produce fertile offspring with them.

The best-documented creations of new species in the laboratory were performed in the late 1980s. William Rice and G.W. Salt bred fruit flies, *Drosophila melanogaster,* using a maze with three different choices of habitat such as light/dark and wet/dry. Each generation was placed into the maze, and the groups of flies that came out of two of the eight exits were set apart to breed with each other in their respective groups. After thirty-five generations, the two groups and their offspring were isolated reproductively because of their strong habitat preferences: they mated only within the areas they preferred, and so did not mate with flies that preferred the other areas. The history of such attempts is described in Rice and Hostert (1993).

Diane Dodd was also able to show how reproductive isolation can develop from mating preferences in *Drosophila pseudoobscura* fruit flies after only eight generations using different food types, starch and maltose.



Dodd's experiment has been easy for many others to replicate, including with other kinds of fruit flies and foods.

Genetics

Few speciation genes have been found. They usually involve the reinforcement process of late stages of speciation. In 2008 a speciation gene causing reproductive isolation was reported. It causes hybrid sterility between related subspecies.

Hybrid speciation

Hybridization between two different species sometimes leads to a distinct phenotype. This phenotype can also be fitter than the parental lineage and as such natural selection may then favor these individuals. Eventually, if reproductive isolation is achieved, it may lead to a separate species. However, reproductive isolation between hybrids and their parents is particularly difficult to achieve and thus hybrid speciation is considered an extremely rare event. The Mariana Mallard is known to have arisen from hybrid speciation.

Hybridization without change in chromosome number is called homoploid hybrid speciation. It is considered very rare but has been shown in *Heliconius* butterflies and sunflowers. Polyploid speciation, which involves changes in chromosome number, is a more common phenomenon, especially in plant species.

Gene transposition as a cause

Theodosius Dobzhansky, who studied fruit flies in the early days of genetic research in 1930s, speculated that parts of chromosomes that switch from one location to another might cause a species to split into two different species. He mapped out how it might be possible for sections of chromosomes to relocate themselves in a genome. Those mobile sections can cause sterility in inter-species hybrids, which can act as a speciation pressure. In theory, his idea was sound, but scientists long debated whether it actually happened in nature. Eventually a competing theory involving the gradual accumulation of mutations was shown to occur in nature so often that geneticists largely dismissed the moving gene hypothesis.

However, 2006 research shows that jumping of a gene from one chromosome to another can contribute to the birth of new species. This validates the reproductive isolation mechanism, a key component of speciation.

Interspersed repeats

Interspersed repetitive DNA sequences function as isolating mechanisms. These repeats protect newly evolving gene sequences from being overwritten by gene conversion, due to the creation of non-homologies between otherwise homologous DNA sequences. The non-homologies create barriers to gene conversion. This barrier allows nascent novel genes to evolve without being overwritten by the progenitors of these genes. This uncoupling allows the evolution of new genes, both within gene families and also allelic forms of a gene. The importance is that this allows the splitting of a gene pool without requiring physical isolation of the organisms harboring those gene sequences.

Human speciation

Humans have genetic similarities with chimpanzees and gorillas, suggesting common ancestors. Analysis of genetic drift and recombination using a Markov model suggests humans and chimpanzees speciated apart 4.1 million years ago.

Chapter- 5

Phenotype

Individuals in the mollusk species *Donax variabilis* show diverse coloration and patterning in their phenotypes.

A **phenotype** is any *observable characteristic* or trait of an organism: such as its morphology, development, biochemical or physiological properties, behavior, and products of behavior (such as a bird's nest). Phenotypes result from the expression of an organism's genes as well as the influence of environmental factors and the interactions between the two.

The genotype of an organism is the inherited instructions it carries within its genetic code. Not all organisms with the same genotype look or act the same way because appearance and behavior are modified by environmental and developmental conditions. Similarly, not all organisms that look alike necessarily have the same genotype.

This genotype-phenotype distinction was proposed by Wilhelm Johannsen in 1911 to make clear the difference between an organism's heredity and what that heredity produces. The distinction is similar to that proposed by August Weismann, who distinguished between germ plasm (heredity) and somatic cells (the body). A more modern version is Francis Crick's Central dogma of molecular biology.

Difficulties in definition

Despite its seemingly straightforward definition, the concept of the phenotype has some hidden subtleties. First, most of the molecules and structures coded by the genetic material are not visible in the appearance of an organism, yet they are observable (for example by Western blotting) and are thus part of the phenotype. Human blood groups are an example. So, by extension, the term phenotype must include characteristics that can be made visible by some technical procedure. Another extension adds behaviour to the phenotype since behaviours are also observable characteristics. Indeed there is research into the clinical relevance of behavioural phenotypes as they pertain to a range of syndromes. Often, the term "phenotype" is incorrectly used as a shorthand to indicate *phenotypical changes* observed in mutated organisms (most often in connection with knockout mice).

Biston betularia morpha *typica*, the standard light-colored Peppered Moth.

Biston betularia morpha *carbonaria*, the melanic Peppered Moth, illustrating discontinuous variation.

Phenotypic variation

Phenotypic variation (due to underlying heritable genetic variation) is a fundamental prerequisite for evolution by natural selection. It is the living organism as a whole that contributes (or not) to the next generation, so natural selection affects the genetic structure of a population indirectly via the contribution of phenotypes. Without phenotypic variation, there would be no evolution by natural selection.

The interaction between genotype and phenotype has often been conceptualized by the following relationship:

genotype + environment → phenotype

A slightly more nuanced version of the relationships is:

genotype + environment + random-variation → phenotype

Genotypes often have much flexibility in the modification and expression of phenotypes; in many organisms these phenotypes are very different under varying environmental conditions. The plant *Hieracium umbellatum* is found growing in two different habitats in Sweden. One habitat is rocky, sea-side cliffs, where the plants are bushy with broad leaves and expanded inflorescences; the other is among sand dunes where the plants grow prostrate with narrow leaves and compact inflorescences. These habitats alternate along the coast of Sweden and the habitat that the seeds of *Hieracium umbellatum* land in, determine the phenotype that grows.

An example of random variation in *Drosophila* flies is the number of ommatidia, which may vary (randomly) between left and right eyes in a single individual as much as they do between different genotypes overall, or between clones raised in different environments.

The concept of phenotype can be extended to variations below the level of the gene that affect an organism's fitness. For example, silent mutations that do not change the corresponding amino acid sequence of a gene may change the frequency of guanine-cytosine base pairs (GC content). These base pairs have a higher thermal stability than adenine-thymine, a property that might convey, among organisms living in high-temperature environments, a selective advantage on variants enriched in GC content.

The Extended Phenotype

The idea of the phenotype has been generalized by Richard Dawkins in *The Extended Phenotype* to mean all the effects a gene has on the outside world that may influence its chances of being replicated. These can be effects on the organism in which the gene resides, the environment, or other organisms.

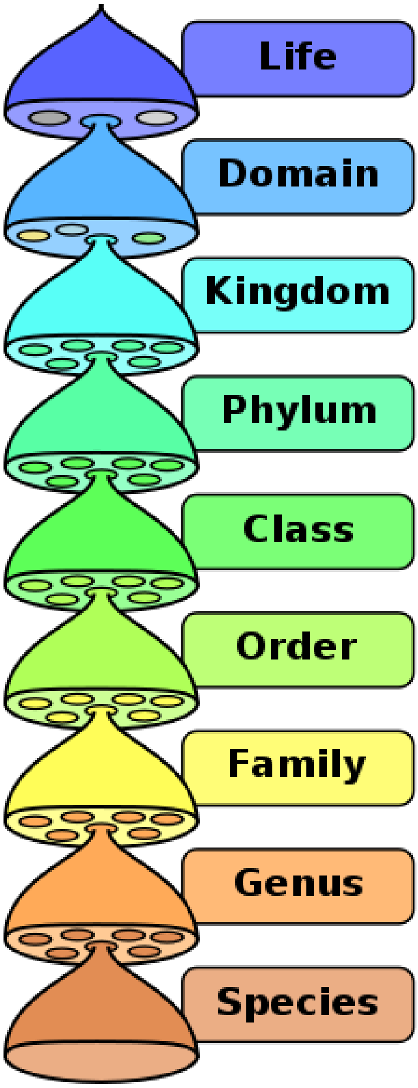
For instance, a beaver dam might be considered a phenotype of beaver genes, the same way beavers' powerful incisor teeth are phenotype expressions of their genes. Dawkins also cites the effect of an organism on the behaviour of another organism (such as the devoted nurturing of a cuckoo by a parent clearly of a different species) as an example of the extended phenotype.

Phenome and phenomics

Although a phenotype is the ensemble of observable characteristics displayed by an organism, the word *phenome* is sometimes used to refer to a collection of traits and their simultaneous study as *phenomics*.

Chapter- 6

Species



The hierarchy of biological classification's eight major taxonomic ranks, which is an example of definition by genus and differentia. A genus contains one or more species. Intermediate minor rankings are not shown.

In biology, a **species** is one of the basic units of biological classification and a taxonomic rank. A species is often defined as a group of organisms capable of interbreeding and producing fertile offspring. While in many cases this definition is adequate, more precise or differing measures are often used, such as similarity of DNA, morphology or ecological niche. Presence of specific locally adapted traits may further subdivide species into subspecies.

The commonly used names for plant and animal taxa sometimes correspond to species: for example, "lion," "walrus," and "Camphor tree" – each refers to a species. In other cases common names do not: for example, "deer" refers to a family of 34 species, including Eld's Deer, Red Deer and Elk (Wapiti). The last two species were once considered a single species, illustrating how species boundaries may change with increased scientific knowledge.

Each species is placed within a single genus. This is a hypothesis that the species is more closely related to other species within its genus than to species of other genera. All species are given a binomial name consisting of the generic name and specific name (or specific epithet). For example, *Boa constrictor*, which is commonly called by its bionomial name, and is one of five species of the *Boa* genus.

A usable definition of the word "species" and reliable methods of identifying particular species are essential for stating and testing biological theories and for measuring biodiversity. Traditionally, multiple examples of a proposed species must be studied for unifying characters before it can be regarded as a species. Extinct species known only from fossils are generally difficult to assign precise taxonomic rankings.

Because of the difficulties with both defining and tallying the total numbers of different species in the world, it is estimated that there are anywhere between 2 and 100 million different species.

Biologists' working definition

A usable definition of the word "species" and reliable methods of identifying particular species is essential for stating and testing biological theories and for measuring biodiversity. Traditionally, multiple examples of a proposed species must be studied for unifying characters before it can be regarded as a species. It is generally difficult to give precise taxonomic rankings to extinct species known only from fossils.

Some biologists may view species as statistical phenomena, as opposed to the traditional idea, with a species seen as a class of organisms. In that case, a species is defined as a separately evolving lineage that forms a single gene pool. Although properties such as DNA-sequences and morphology are used to help separate closely related lineages, this definition has fuzzy boundaries. However, the exact definition of the term "species" is still controversial, particularly in prokaryotes, and this is called the species problem. Biologists have proposed a range of more precise definitions, but the definition used is a pragmatic choice that depends on the particularities of the species of concern.

Common names and species

The commonly used names for plant and animal taxa sometimes correspond to species: for example, "lion", "walrus", and "Camphor tree" – each refers to a species. In other cases common names do not: for example, "deer" refers to a family of 34 species, including Eld's Deer, Red Deer and Elk (Wapiti). The last two species were once considered a single species, illustrating how species boundaries may change with increased scientific knowledge.

Because of the difficulties with both defining and tallying the total numbers of different species in the world, it is estimated that there are anywhere between 2 and 100 million different species.

Placement within genera

Ideally, a species is given a formal, scientific name, although in practice there are very many unnamed species (which have only been described, not named). When a species is named, it is placed within a genus. From a scientific point of view this can be regarded as a hypothesis that the species is more closely related to other species within its genus (if any) than to species of other genera. Species and genus are usually defined as part of a larger taxonomic hierarchy. The best-known taxonomic ranks are, in order: life, domain, kingdom, phylum, class, order, family, genus, and species. This assignment to a genus is not immutable; later a different (or the same) taxonomist may assign it to a different genus, in which case the name will also change.

In biological nomenclature, the name for a species is a two-part name (a binomial name), treated as Latin, although roots from any language can be used as well as names of locales or individuals. The generic name is listed first (with its leading letter capitalized), followed by a second term, the specific name (or specific epithet). For example, the species commonly known as the Longleaf Pine is *Pinus palustris*; gray wolves belong to the species *Canis lupus*, coyotes to *Canis latrans*, golden jackals to *Canis aureus*, etc., and all of those belong to the genus *Canis* (which also contains many other species). The name of the species is the whole binomial, not just the second term (which may be called the specific name for animals).

This binomial naming convention, later formalized in the biological codes of nomenclature, was first used by Leonhart Fuchs and introduced as the standard by Carolus Linnaeus in his 1753, *Species Plantarum* (followed by his, 1758 *Systema Naturae*, 10th edition). At that time, the chief biological theory was that species represented independent acts of creation by God and were therefore considered objectively real and immutable, so the hypothesis of common descent did not apply.

Abbreviated names

Books and articles sometimes intentionally do not identify species fully and use the abbreviation "**sp.**" in the singular or "**spp.**" in the plural in place of the specific epithet: for example, ***Canis*** **sp.** This commonly occurs in the following types of situations:

- The authors are confident that some individuals belong to a particular genus but are not sure to which exact species they belong. This is particularly common in paleontology.

- The authors use "spp." as a short way of saying that something applies to many species within a genus, but do not wish to say that it applies to all species within that genus. If scientists mean that something applies to all species within a genus, they use the genus name without the specific epithet.

In books and articles, genus and species names are usually printed in italics. If using "sp." and "spp.", these should not be italicized.

Difficulty of defining "species" and identifying particular species

The Greenish Warbler demonstrates the concept of a ring species.

It is surprisingly difficult to define the word "species" in a way that applies to all naturally occurring organisms, and the debate among biologists about how to define "species" and how to identify actual species is called the species problem. Over two dozen distinct definitions of "species" are in use amongst biologists.

Most textbooks follow Ernst Mayr's definition of a species as "groups of actually or potentially interbreeding natural populations, which are reproductively isolated from other such groups".

Various parts of this definition serve to exclude some unusual or artificial matings:

- Those that occur only in captivity (when the animal's normal mating partners may not be available) or as a result of deliberate human action
- Animals that may be physically and physiologically capable of mating but, for various reasons, do not normally do so in the wild

The typical textbook definition above works well for most multi-celled organisms, but there are several types of situations in which it breaks down:

- By definition it applies only to organisms that reproduce sexually. So it does not work for asexually reproducing single-celled organisms and for the relatively few parthenogenetic multi-celled organisms. The term "phylotype" is often applied to such organisms.
- Biologists frequently do not know whether two morphologically similar groups of organisms are "potentially" capable of interbreeding.
- There is considerable variation in the degree to which hybridization may succeed under natural conditions, or even in the degree to which some organisms use sexual reproduction between individuals to breed.
- In ring species, members of adjacent populations interbreed successfully but members of some non-adjacent populations do not.
- In a few cases it may be physically impossible for animals that are members of the same species to mate. However, these are cases in which human intervention has caused gross morphological changes, and are therefore excluded by the biological species concept.

Horizontal gene transfer makes it even more difficult to define the word "species". There is strong evidence of horizontal gene transfer between very dissimilar groups of prokaryotes, and at least occasionally between dissimilar groups of eukaryotes; and Williamson argues that there is evidence for it in some crustaceans and echinoderms. All definitions of the word "species" assume that an organism gets all its genes from one or two parents that are very like that organism, but horizontal gene transfer makes that assumption false.

Definitions of species

The question of how best to define "species" is one that has occupied biologists for centuries, and the debate itself has become known as the species problem. Darwin wrote in chapter II of *On the Origin of Species*:

No one definition has satisfied all naturalists; yet every naturalist knows vaguely what he means when he speaks of a species. Generally the term includes the unknown element of a distinct act of creation.

But later, in *The Descent of Man*, when addressing "The question whether mankind consists of one or several species", Darwin revised his opinion to say:

it is a hopeless endeavour to decide this point on sound grounds, until some definition of the term "species" is generally accepted; and the definition must not include an element that cannot possibly be ascertained, such as an act of creation.

The modern theory of evolution depends on a fundamental redefinition of "species". Prior to Darwin, naturalists viewed species as ideal or general types, which could be exemplified by an ideal specimen bearing all the traits general to the species. Darwin's theories shifted attention from uniformity to variation and from the general to the particular. According to intellectual historian Louis Menand,

Once our attention is redirected to the individual, we need another way of making generalizations. We are no longer interested in the conformity of an individual to an ideal type; we are now interested in the relation of an individual to the other individuals with which it interacts. To generalize about groups of interacting individuals, we need to drop the language of types and essences, which is prescriptive (telling us what finches should be), and adopt the language of statistics and probability, which is predictive (telling us what the average finch, under specified conditions, is likely to do). Relations will be more important than categories; functions, which are variable, will be more important than purposes; transitions will be more important than boundaries; sequences will be more important than hierarchies.

This shift results in a new approach to "species"; Darwin

concluded that species are what they appear to be: ideas, which are provisionally useful for naming groups of interacting individuals. "I look at the term species", he wrote, "as one arbitrarily given for the sake of convenience to a set of individuals closely resembling each other ... It does not essentially differ from the word variety, which is given to less distinct and more fluctuating forms. The term variety, again, in comparison with mere individual differences, is also applied arbitrarily, and for convenience sake."

Practically, biologists define species as *populations of organisms that have a high level of genetic similarity*. This may reflect an adaptation to the same niche, and the transfer of genetic material from one individual to others, through a variety of possible means. The exact level of similarity used in such a definition is arbitrary, but this is the most common definition used for organisms that reproduce asexually (asexual reproduction), such as some plants and microorganisms.

This lack of any clear species concept in microbiology has led to some authors arguing that the term "species" is not useful when studying bacterial evolution. Instead they see genes as moving freely between even distantly related bacteria, with the entire bacterial domain being a single gene pool. Nevertheless, a kind of rule of thumb has been established, saying that species of *Bacteria* or *Archaea* with 16S rRNA gene sequences more similar than 97% to each other need to be checked by DNA-DNA Hybridization if they belong to the same species or not. This concept has been updated recently, saying that the border of 97% was too low and can be raised to 98.7%.

In the study of sexually reproducing organisms, where genetic material is shared through the process of reproduction, the ability of two organisms to interbreed and produce fertile offspring of both sexes is generally accepted as a simple indicator that the organisms share enough genes to be considered members of the same species. Thus a "species" is a group of interbreeding organisms.

This definition can be extended to say that a species is a group of organisms that could potentially interbreed – fish could still be classed as the same species even if they live in different lakes, as long as they could still interbreed were they ever to come into contact with each other. On the other hand, there are many examples of series of three or more distinct populations, where individuals of the population in the middle can interbreed with the populations to either side, but individuals of the populations on either side cannot interbreed. Thus, one could argue that these populations constitute a single species, or two distinct species. This is not a paradox; it is evidence that species are defined by gene frequencies, and thus have fuzzy boundaries.

Consequently, any single, universal definition of "species" is necessarily arbitrary. Instead, biologists have proposed a range of definitions; which definition a biologists uses is a pragmatic choice, depending on the particularities of that biologist's research.

Typological species

> A group of organisms in which individuals are members of the species if they sufficiently conform to certain fixed properties or "rights of passage". The clusters of variations or phenotypes within specimens (i.e. longer or shorter tails) would differentiate the species. This method was used as a "classical" method of determining species, such as with Linnaeus early in evolutionary theory. However, we now know that different phenotypes do not always constitute different species (e.g.: a 4-winged Drosophila born to a 2-winged mother is not a different species). Species named in this manner are called *morphospecies*

Morphological species

> A population or group of populations that differs morphologically from other populations. For example, we can distinguish between a chicken and a duck because they have different shaped bills and the duck has webbed feet. Species have been defined in this way since well before the beginning of recorded history. This species concept is highly criticized because more recent genetic data reveal that genetically distinct populations may look very similar and, contrarily, large morphological differences sometimes exist between very closely related populations. Nonetheless, most species known have been described solely from morphology.

Biological / Isolation species

> A set of actually or potentially interbreeding populations. This is generally a useful formulation for scientists working with living examples of the higher taxa like mammals, fish, and birds, but more problematic for organisms that do not reproduce sexually. The results of breeding experiments done in artificial conditions may or may not reflect what would happen if the same organisms

encountered each other in the wild, making it difficult to gauge whether or not the results of such experiments are meaningful in reference to natural populations.

Biological / reproductive species

Two organisms that are able to reproduce naturally to produce fertile offspring of both sexes. Organisms that can reproduce but almost always make infertile hybrids of at least one sex, such as a mule, hinny or F1 male cattalo are not considered to be the same species.

Recognition species

Based on shared reproductive systems, including mating behavior. The Recognition concept of species has been introduced by Hugh E. H. Paterson.

Mate-recognition species

A group of organisms that are known to recognize one another as potential mates. Like the isolation species concept above, it applies only to organisms that reproduce sexually. Unlike the isolation species concept, it focuses specifically on pre-mating reproductive isolation.

Evolutionary / Darwinian species

A group of organisms that shares an ancestor; a lineage that maintains its integrity with respect to other lineages through both time and space. At some point in the progress of such a group, some members may diverge from the main population and evolve into a subspecies, a process that eventually will lead to the formation of a new full species if isolation (geographical or ecological) is maintained.

Phylogenetic (Cladistic)

A group of organisms that shares an ancestor; a lineage that maintains its integrity with respect to other lineages through both time and space. At some point in the progress of such a group, members may diverge from one another: when such a divergence becomes sufficiently clear, the two populations are regarded as separate species. This differs from evolutionary species in that the parent species goes extinct taxonomically when a new species evolve, the mother and daughter populations now forming two new species. Subspecies as such are not recognized under this approach; either a population is a phylogenetic species or it is not taxonomically distinguishable.

Ecological species

A set of organisms adapted to a particular set of resources, called a niche, in the environment. According to this concept, populations form the discrete phonetic clusters that we recognize as species because the ecological and evolutionary processes controlling how resources are divided up tend to produce those clusters.

Genetic species

Based on similarity of DNA of individuals or populations. Techniques to compare similarity of DNA include DNA-DNA hybridization, and genetic fingerprinting (or DNA barcoding).

Phenetic species

Based on phenotypes.

Microspecies

Species that reproduce without meiosis or fertilization so that each generation is genetically identical to the previous generation.

Cohesion species

Most inclusive population of individuals having the potential for phenotypic cohesion through intrinsic cohesion mechanisms. This is an expansion of the mate-recognition species concept to allow for post-mating isolation mechanisms; no matter whether populations can hybridize successfully, they are still distinct cohesion species if the amount of hybridization is insufficient to completely mix their respective gene pools.

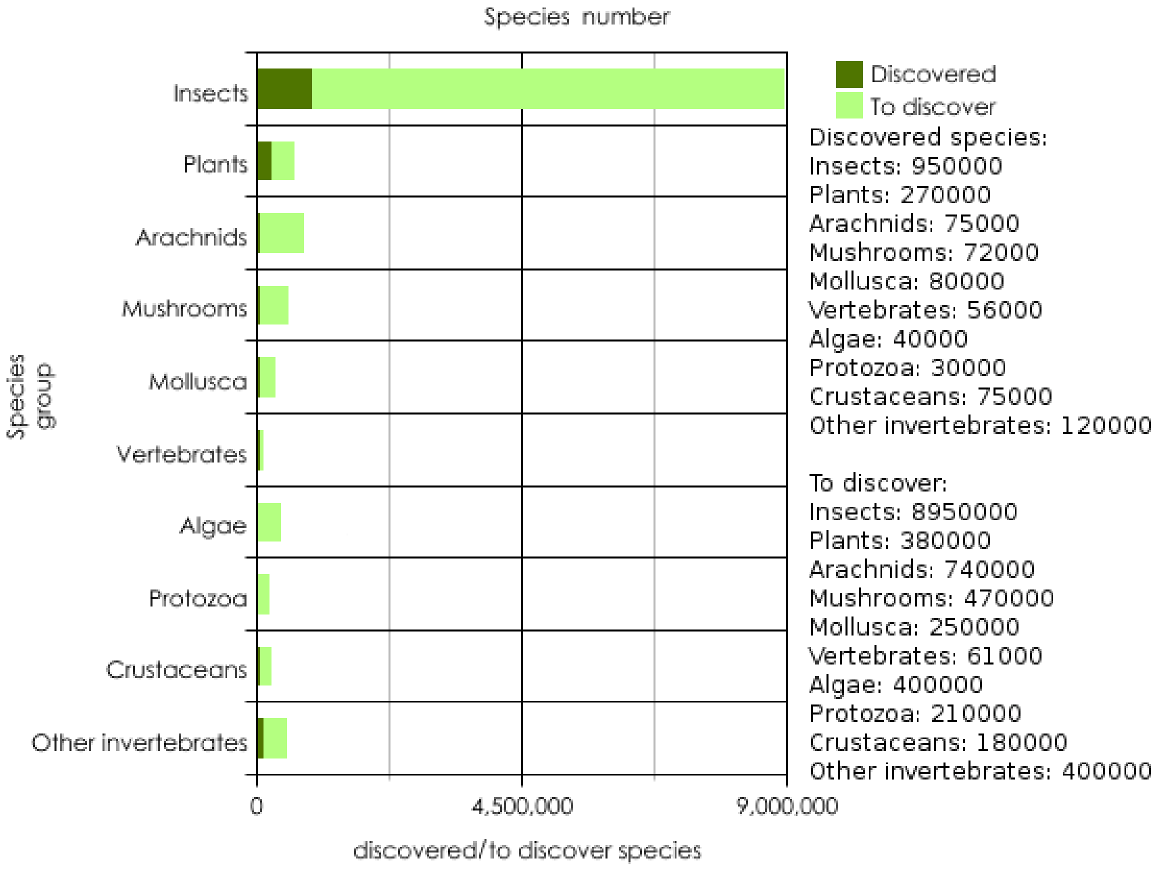
Evolutionarily Significant Unit (ESU)

An evolutionarily significant unit is a population of organisms that is considered distinct for purposes of conservation. Often referred to as a species or a *wildlife species*, an ESU also has several possible definitions, which coincide with definitions of species.

In practice, these definitions often coincide, and the differences between them are more a matter of emphasis than of outright contradiction. Nevertheless, no species concept yet proposed is entirely objective, or can be applied in all cases without resorting to judgment. Given the complexity of life, some have argued that such an objective definition is in all likelihood impossible, and biologists should settle for the most practical definition.

For most vertebrates, this is the biological species concept (BSC), and to a lesser extent (or for different purposes) the phylogenetic species concept (PSC). Many BSC subspecies are considered species under the PSC; the difference between the BSC and the PSC can be summed up insofar as that the BSC defines a species as a consequence of manifest evolutionary *history*, while the PSC defines a species as a consequence of manifest evolutionary *potential*. Thus, a PSC species is "made" as soon as an evolutionary lineage has started to separate, while a BSC species starts to exist only when the lineage separation is complete. Accordingly, there can be considerable conflict between alternative classifications based upon the PSC versus BSC, as they differ completely in their treatment of taxa that would be considered subspecies under the latter model (e.g., the numerous subspecies of honey bees).

Numbers of species



Undiscovered and discovered species

Bearing in mind the aforementioned problems with categorising species, the following numbers are only a soft guide. In 2007, they broke down as follows:

Total number of species (estimated): 7–100 millions (identified and unidentified), including:

- 5–10 million bacteria;
- 74,000–120,000 fungi;

Of the *identified* eukaryote species we have:

- 1.6 million, including:
    - 297,326 plants, including:
        - 15,000 mosses,
        - 13,025 Ferns and horsetails,
        - 980 gymnosperms,
        - 258,650 angiosperms,

        - 199,350 dicotyledons,
        - 59,300 monocotyledons,
    - 9,671 Red and green algae,
- 28,849 fungi & other non-animals, including:
    - 10,000 lichens,
    - 16,000 mushrooms,
    - 2,849 brown algae,
- 1,250,000 animals, including:
    - 1,203,375 invertebrates:
        - 950,000 insects,
        - 81,000 mollusks,
        - 40,000 crustaceans,
        - 2,175 corals,
        - 130,200 others;
    - 59,811 vertebrates:
        - 29,300 fish,
        - 6,199 amphibians,
        - 8,240 reptiles,
        - 9,956 birds,
        - 5,416 mammals

At present, organisations such as the Global Taxonomy Initiative, the European Distributed Institute of Taxonomy and the Census of Marine Life (the latter only for marine organisms) are trying to improve taxonomy and implement previously undiscovered species to the taxonomy system. Because we know but a portion of the organisms in the biosphere, we do not have a complete understanding of the workings of our environment. To make matters worse, despite the discovery of new species, according to professor James Mallet, we are wiping out these species at an unprecedented rate. This means that even before a new species has had the chance of being studied and classified, it may already be extinct.

Importance in biological classification

The idea of *species* has a long history. It is one of the most important levels of classification, for several reasons:

- It often corresponds to what lay people treat as the different basic kinds of organism – dogs are one species, cats another.
- It is the standard binomial nomenclature (or trinomial nomenclature) by which scientists typically refer to organisms.
- It is the highest taxonomic level that cannot be made more or less inclusionary.

After years of use, the concept remains central to biology and a host of related fields, and yet also remains at times ill-defined.

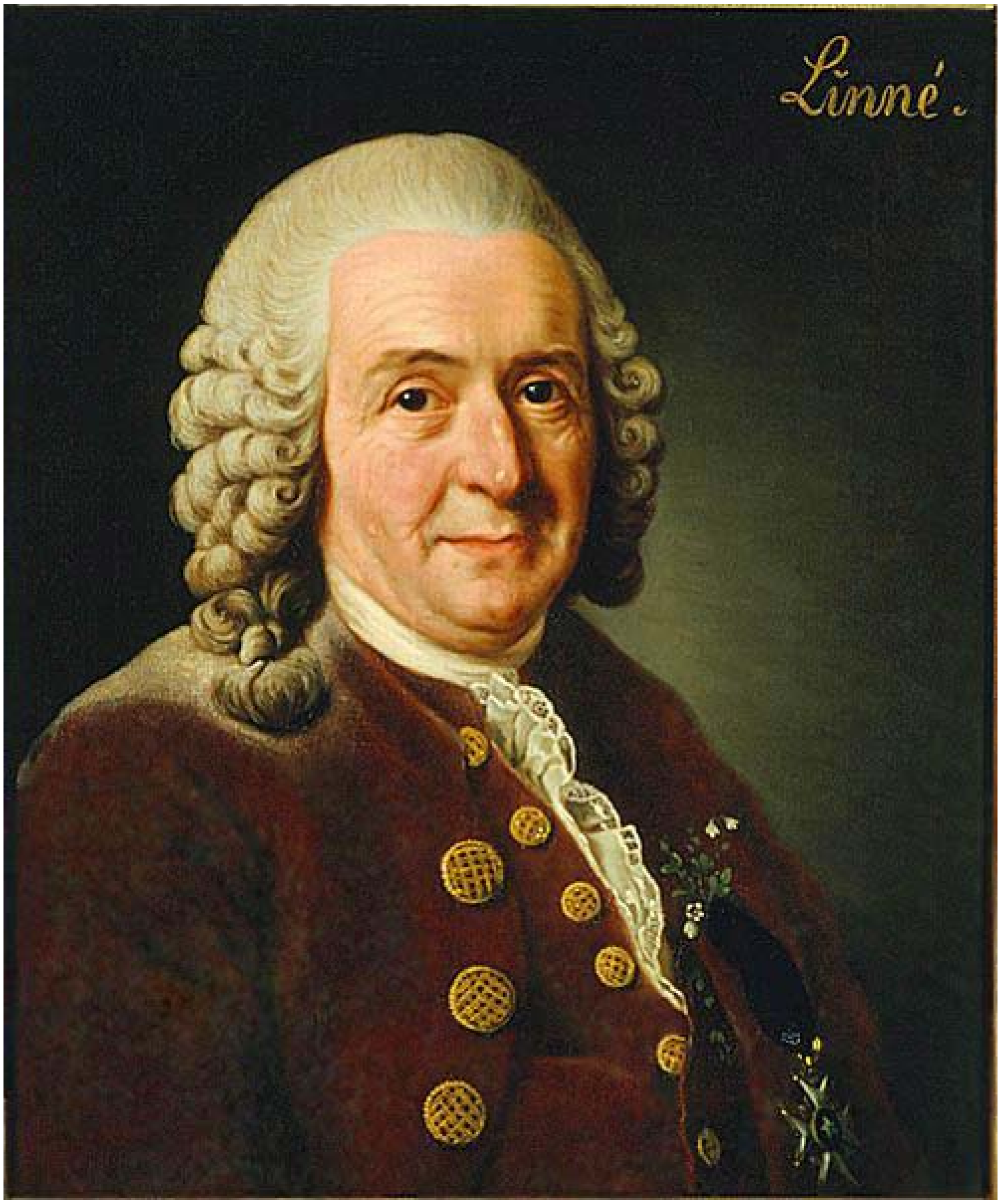
Implications of assignment of species status

The naming of a particular species may be regarded as a *hypothesis* about the evolutionary relationships and distinguishability of that group of organisms. As further information comes to hand, the hypothesis may be confirmed or refuted. Sometimes, especially in the past when communication was more difficult, taxonomists working in isolation have given two distinct names to individual organisms later identified as the same species. When two named species are discovered to be of the same species, the older species name is usually retained, and the newer species name dropped, a process called *synonymization*, or colloquially, as **lumping**. Dividing a taxon into multiple, often new, taxons is called **splitting**. Taxonomists are often referred to as "lumpers" or "splitters" by their colleagues, depending on their personal approach to recognizing differences or commonalities between organisms.

Traditionally, researchers relied on observations of anatomical differences, and on observations of whether different populations were able to interbreed successfully, to distinguish species; both anatomy and breeding behavior are still important to assigning species status. As a result of the revolutionary (and still ongoing) advance in microbiological research techniques, including DNA analysis, in the last few decades, a great deal of additional knowledge about the differences and similarities between species has become available. Many populations formerly regarded as separate species are now considered a single taxon, and many formerly grouped populations have been split. Any taxonomic level (species, genus, family, etc.) can be synonymized or split, and at higher taxonomic levels, these revisions have been still more profound.

From a taxonomical point of view, groups within a species can be defined as being of a taxon hierarchically lower than a species. In zoology only the subspecies is used, while in botany the variety, subvariety, and form are used as well. In conservation biology, the concept of evolutionary significant units (ESU) is used, which may be define either species or smaller distinct population segments. Identifying and naming species is the providence of alpha taxonomy.

Historical development of the species concept



Linnaeus believed in the fixity of species.

In the earliest works of science, a species was simply an individual organism that represented a group of similar or nearly identical organisms. No other relationships beyond that group were implied. Aristotle used the words *genus* and *species* to mean generic and specific categories. Aristotle and other pre-Darwinian scientists took the species to be distinct and unchanging, with an "essence", like the chemical elements. When early observers began to develop systems of organization for living things, they began to place formerly isolated species into a context. Many of these early delineation schemes would now be considered whimsical and these included consanguinity based on color (all plants with yellow flowers) or behavior (snakes, scorpions and certain biting ants).

In the 18th century Swedish scientist Carolus Linnaeus classified organisms according to differences in the form of reproductive apparatus. Although his system of classification sorts organisms according to degrees of similarity, it made no claims about the relationship between similar species. At that time, it was still widely believed that there was no organic connection between species, no matter how similar they appeared. This approach also suggested a type of idealism: the notion that each species existed as an "ideal form". Although there are always differences (although sometimes minute) between individual organisms, Linnaeus considered such variation problematic. He strove to identify individual organisms that were exemplary of the species, and considered other non-exemplary organisms to be deviant and imperfect.

By the 19th century most naturalists understood that species could change form over time, and that the history of the planet provided enough time for major changes. Jean-Baptiste Lamarck, in his 1809 *Zoological Philosophy*, offered one of the first logical arguments against creationism. The new emphasis was on determining *how* a species could change over time. Lamarck suggested that an organism could pass on an acquired trait to its offspring, i.e., the giraffe's long neck was attributed to generations of giraffes stretching to reach the leaves of higher treetops (this well-known and simplistic example, however, does not do justice to the breadth and subtlety of Lamarck's ideas). With the acceptance of the natural selection idea of Charles Darwin in the 1860s, however, Lamarck's view of goal-oriented evolution, also known as a teleological process, was eclipsed. Recent interest in inheritance of acquired characteristics centers around epigenetic processes, e.g. methylation, that do not affect DNA sequences, but instead alter expression in an inheritable manner. Thus, neo-lamarckism, as it is sometimes termed, is not a challenge to the theory of evolution by natural selection.

Charles Darwin and Alfred Wallace provided what scientists now consider as the most powerful and compelling theory of evolution. Darwin argued that it was populations that evolved, not individuals. His argument relied on a radical shift in perspective from that of Linnaeus: rather than defining species in ideal terms (and searching for an ideal representative and rejecting deviations), Darwin considered variation among individuals to be natural. He further argued that variation, far from being problematic, actually provides the *explanation* for the existence of distinct species.

Darwin's work drew on Thomas Malthus' insight that the rate of growth of a biological population will always outpace the rate of growth of the resources in the environment, such as the food supply. As a result, Darwin argued, not all the members of a population will be able to survive and reproduce. Those that did will, on average, be the ones possessing variations—however slight—that make them slightly better adapted to the environment. If these variable traits are heritable, then the offspring of the survivors will also possess them. Thus, over many generations, adaptive variations will accumulate in the population, while counter-adaptive traits will tend to be eliminated.

It should be emphasized that whether a variation is adaptive or non-adaptive depends on the environment: different environments favor different traits. Since the environment effectively selects which organisms live to reproduce, it is the environment (the "fight for existence") that selects the traits to be passed on. This is the theory of evolution by natural selection. In this model, the length of a giraffe's neck would be explained by positing that proto-giraffes with longer necks would have had a significant reproductive advantage to those with shorter necks. Over many generations, the entire population would be a species of long-necked animals.

In 1859, when Darwin published his theory of natural selection, the mechanism behind the inheritance of individual traits was unknown. Although Darwin made some speculations on how traits are inherited (pangenesis), his theory relies only on the fact that inheritable traits *exist*, and are variable (which makes his accomplishment even more remarkable.) Although Gregor Mendel's paper on genetics was published in 1866, its significance was not recognized. It was not until 1900 that his work was rediscovered by Hugo de Vries, Carl Correns and Erich von Tschermak, who realised that the "inheritable traits" in Darwin's theory are genes.

The theory of the evolution of species through natural selection has two important implications for discussions of species—consequences that fundamentally challenge the assumptions behind Linnaeus' taxonomy. First, it suggests that species are not just similar, they may actually be related. Some students of Darwin argue that *all* species are descended from a common ancestor. Second, it supposes that "species" are not homogeneous, fixed, permanent things; members of a species are all different, and over time species change. This suggests that species do not have any clear boundaries but are rather momentary statistical effects of constantly changing gene-frequencies. One may still use Linnaeus' taxonomy to identify individual plants and animals, but one can no longer think of species as independent and immutable.

The rise of a new species from a parental line is called speciation. There is no clear line demarcating the ancestral species from the descendant species.

Although the current scientific understanding of species suggests that there is no rigorous and comprehensive way to distinguish between different species in *all* cases, biologists continue to seek concrete ways to operationalize the idea. One of the most popular biological definitions of species is in terms of reproductive isolation; if two creatures cannot reproduce to produce fertile offspring of both sexes, then they are in different species. This definition captures a number of intuitive species boundaries, but it remains imperfect. It has nothing to say about species that reproduce asexually, for example, and it is very difficult to apply to extinct species. Moreover, boundaries between species are often fuzzy: there are examples where members of one population can produce fertile offspring of both sexes with a second population, and members of the second population can produce fertile offspring of both sexes with members of a third population, but members of the first and third population cannot produce fertile offspring, or can only produce fertile offspring of the homozygous sex. Consequently, some people reject this definition of a species.

Richard Dawkins defines two organisms as conspecific if and only if they have the same number of chromosomes and, for each chromosome, both organisms have the same number of nucleotides (*The Blind Watchmaker*, p. 118). However, most if not all taxonomists would strongly disagree. For example, in many amphibians, most notably in New Zealand's *Leiopelma* frogs, the genome consists of "core" chromosomes that are mostly invariable and accessory chromosomes, of which exist a number of possible combinations. Even though the chromosome numbers are highly variable between populations, these can interbreed successfully and form a single evolutionary unit. In plants, polyploidy is extremely commonplace with few restrictions on interbreeding; as individuals with an odd number of chromosome sets are usually sterile, depending on the actual number of chromosome sets present, this results in the odd situation where some individuals of the same evolutionary unit can interbreed with certain others and some cannot, with all populations being eventually linked as to form a common gene pool.

The classification of species has been profoundly affected by technological advances that have allowed researchers to determine relatedness based on molecular markers, starting with the comparatively crude blood plasma precipitation assays in the mid-20th century to Charles Sibley's ground-breaking DNA-DNA hybridization studies in the 1970s leading to DNA sequencing techniques. The results of these techniques caused revolutionary changes in the higher taxonomic categories (such as phyla and classes), resulting in the reordering of many branches of the phylogenetic tree. For taxonomic categories below genera, the results have been mixed so far; the pace of evolutionary change on the molecular level is rather slow, yielding clear differences only after considerable periods of reproductive separation. DNA-DNA hybridization results have led to misleading conclusions, the Pomarine Skua – Great Skua phenomenon being a famous example. Turtles have been determined to evolve with just one-eighth of the speed of other reptiles on the molecular level, and the rate of molecular evolution in albatrosses is half of what is found in the rather closely related storm-petrels. The hybridization technique is now obsolete and is replaced by more reliable computational approaches for sequence comparison. Molecular taxonomy is not directly based on the evolutionary processes, but rather on the overall change brought upon by these processes. The processes that lead to the generation and maintenance of variation such as mutation, crossover and selection are not uniform. DNA is only extremely rarely a direct target of natural selection rather than changes in the DNA sequence enduring over generations being a result of the latter; for example, silent transition-transversion combinations would alter the melting point of the DNA sequence, but not the sequence of the encoded proteins and thus are a possible example where, for example in microorganisms, a mutation confers a change in fitness all by itself.

Species as taxa

The scientific name of a species (often of Latin or Greek origin) in the binominal nomenclature introduced by Carl Linnaeus in 1753 is composed of two parts, which are written in italic font and for which there are different expressions in botany and zoology. The first part of that name is spelled upper case and is in both disciplines known as the genus name (also called the *generic name*). The second part is always spelled lower case and in botany it is called the *specific epithet*.

- Example: in the European beech (*Fagus sylvatica*) the component *Fagus* refers to the genus, *sylvatica* is the specific epithet.

In zoology, the second part is called the *specific name*.

- Example: in the lion (*Panthera leo*) the component *Panthera* refers to the genus, *leo* is the specific name.

The scientific name is completed when authors, years and parentheses are added.

In botany the name of the author is usually abbreviated, for example "L." stands for "Linnaeus".

- Example: shiitake *Lentinula edodes* (Berk.) Pegler
  M. J. Berkeley was the first to describe the species, D. Pegler has placed it into the currently used system.

In zoology, the addition of author and year are optional, *Panthera leo* is thus an entirely correct name. The International Code of Zoological Nomenclature prescribes how to add author(s) (if possible not abbreviated) and year (or author alone without year). If the species is cited in different genus than the one in which it was originally described, author(s) and year are given in parentheses. Between author and year often a comma is set (but not required).

- Example: lion *Panthera leo* (Linnaeus, 1758)
  Carl Nilsson Linnæus described the lion first and as *Felis leo*. The person who first placed the lion in the genus *Panthera* Oken, 1816 is not relevant in zoology. Instead of Linnæus, usually Linnaeus is spelled.

Chapter- 7

Species Problem

> “ ... I was much struck how entirely vague and arbitrary is the distinction between species and varieties ”
>
> — Charles Darwin, *On the Origin of Species* (p. 48)

The **species problem** is a mixture of difficult, related questions that often come up when biologists identify species and when they define the word "species".

One common but sometimes difficult question is how best to decide just which particular species an organism belongs to. Another challenge is deciding when to recognize a new species. This is a question for the biologist who discovers organisms that appear to be different from those that belong to already described species. A related question arises when new data indicate that one previously described species actually may include two or more separately evolving groups, each of which could possibly be recognized as a separate species.

Many of the debates on species touch on philosophical issues, such as nominalism and realism, as well as on issues of language and cognition.

This current meaning of the phrase "species problem" is quite different from what was meant by "species problem" during the 19th and early 20th centuries, as used by Darwin and others. For Darwin the species problem was the question of how new species arose.

Confusion on the meaning of "*Species*"

Species is one of several ranks in the hierarchical system of scientific classification. These are called *taxonomic ranks*, and the system of classification includes, in addition to *species* the ranks of genus and family and others all the way up to kingdom. Usually the rank of species is the *basal rank*, meaning that in the system of scientific classification *species* is the bottommost rank that includes no other ranks. However sometimes when one species, that is already named and described, is found to actually include two slightly different kinds of organisms, it is necessary to use the rank of *subspecies*.

Even though it is not disputed that *species* is a taxonomic rank, this does not prevent disagreements when particular species are discussed. Consider the case of the Baltimore oriole and Bullock's oriole, two similar species of birds that have sometimes in the past been considered to be one single species. Currently biologists agree that these are actually two separate species, but in the past this was not the case.

It is common in debates about species for participants to argue at cross purposes. For example, in a debate over the species status of Baltimore Oriole and Bullock's Oriole one person might think that the critical question is about the two kinds of orioles and how similar they are. A second person might think that the critical question concerns the actual taxonomic rank of species, and on what the correct criteria are for identifying a species. If one person is talking about the birds, and another person is talking about the rank of species, then there can be confusion.

Disagreements and confusion also happen over just what the best criteria are for identifying new species. In 1942 the famous biologist Ernst Mayr wrote that because biologists have different ways of identifying species, they actually have different species concepts. Mayr proceeded to list five different species concepts, and since then many more have been added. The question of which species concept is *best* has occupied many printed pages and many hours of discussion.

Some debates are philosophical in nature. One common disagreement is over whether a species is defined by the characteristics that biologists use to identify the species, or whether a species is an evolving entity in nature. Every named species has been formally described as a type of organism with particular defining characteristics. These defining traits are used to identify which species organisms belong to. But for many species, all of the individuals that fit the defining criteria also make up a single evolving unit. These two different ways of thinking about species, as a category and as an evolving population, are quite different from each other.

History

Before Darwin

The idea that one organism reproduces by giving birth to a similar organism, or producing seeds that grow to a similar organism, goes back to the earliest days of farming. While people tended to think of this as a relatively stable process, many thought that change was possible. The term *species* was just used as a term for a sort or kind of organism, until in 1686 John Ray introduced the biological concept that species were distinguished by always producing the same species, and this was fixed and permanent, though considerable variation was possible within a species. Carolus Linnaeus (1707–1778) formalized the taxonomic rank of species, and devised the two part naming system of binomial nomenclature that we use today. However this did not prevent disagreements on the best way to identify species.

The history of definitions of the term "species" reveal that the seeds of the modern species debate were alive and growing long before Darwin.

From Darwin to Mayr

Charles Darwin's famous book *On the Origin of Species* (1859) offered an explanation as to how species changed over time (evolution). Although Darwin did not provide details on how one species splits into two, he viewed speciation as a gradual process. If Darwin was correct, then when new *incipient species* are forming there must be a period of time when they are not yet distinct enough to be recognized as species. Darwin's theory suggested that there was often not going to be an objective fact of the matter, on whether there were one or two species.

Darwin's book triggered a crisis of uncertainty for some biologists over the objectivity of species, and some came to wonder whether individual species could be objectively real — i.e. have an existence that is independent of the human observer.

In the 1920s and 1930s, Mendel's theory of inheritance and Darwin's theory of evolution by natural selection were joined in what was called the modern evolutionary synthesis. This conjunction of theories also had a large impact on how biologists think about species. Edward Poulton anticipated many ideas on species that today are well accepted, and that were later more fully developed by Theodosius Dobzhansky and Ernst Mayr, two of the architects of the modern synthesis. Dobzhansky's 1937 book articulated the genetic processes that occur when incipient species are beginning to diverge. In particular Dobzhansky described the critical role, for the formation of new species, of the evolution of reproductive isolation.

Mayr and recent history

Ernst Mayr's 1942 book was a turning point for the species problem. In it he wrote about how different investigators approach species identification, and he characterized these different approaches as different species concepts. He also argued strongly for, what came to be called, a *Biological Species Concept* (BSC), which is that a species consists of populations of organisms that can reproduce with one another and that are reproductively isolated from other such populations.

Mayr was not the first to define "species" on the basis of reproductive compatibility. Many others before Mayr had suggested this idea, as Mayr makes clear in his book on the history of biology. For example Mayr discusses how Buffon proposed this kind of definition of "species" in 1753. The idea of shared reproduction within species is even contained in the Biblical story of Noah's ark, in which each species was preserved by saving a reproductive pair.

Theodosius Dobzhansky was a close contemporary of Mayr's and the author of a classic book, that came out a few years before Mayr's, that was about the evolutionary origins of reproductive barriers between species. Many biologists credit Dobzhansky and Mayr jointly for emphasizing the need to consider reproductive isolation when studying species and speciation.

Mayr was persuasive in many respects and from 1942 until his death in 2005 he and the biological species concept (BSC) played a central role in nearly all debates on the species problem. For many, the Biological Species Concept was a useful theoretical idea because it leads to a focus on the evolutionary origins of barriers to reproduction between species. But the BSC has been criticized for not being very useful for deciding when to identify new species. It is also true that there are many cases where members of different species will hybridize and produce fertile offspring when they are under confined conditions, such as in zoos. One fairly extreme example is that lions and tigers will hybridize in captivity, and at least some of the offspring have been reported to be fertile. Mayr's response to cases like these is that the reproductive barriers that are important for species are the ones that occur in the wild. But even so it is also the case that there are many cases of different species that are known to hybridize and produce fertile offspring in nature.

After Mayr's 1942 book many more species concepts were introduced. Some, such as the Phylogenetic Species Concept (PSC), were designed to be more useful than the BSC for actually deciding when a new species should be described. However not all of the new species concepts were about identifying species, and some concepts were mostly conceptual or philosophical.

About two dozen species concepts have been identified or proposed since Mayr's 1942 book, and many articles and several books have been written on the species problem. At some point it became common for articles to profess to "solve" or "dissolve" the species problem.

Some have argued that the species problem is too multidimensional to be "solved" by one definition of species or one species concept. Since the 1990s articles have appeared that make the case that species concepts, particularly those that specify how species should be identified, have not been very helpful in resolving the species problem.

Although Mayr promoted the Biological Species Concept for use in systematics, the concept has been criticized as not being useful for those who do research in systematics. Some systematists have criticized the BSC as not being operational. However for many others the BSC is the preferred description of species. For example many geneticists who work on the process of species formation prefer the BSC because it emphasizes the role of barriers to reproduction between species.

Philosophical aspects

Realism and nominalism

Realism and Nominalism are philosophical subjects that come up in debates over whether or not species literally exist. From one perspective, each species is a kind of organism and each species is based on a set of characteristics that are shared by all the organisms in the species. This usage of "species" refers to the taxonomic sense of the word, and under this kind of meaning a species is a category, or a type, or a *natural kind*. For example, the species that we call *giraffe* is a category of things that people have recognized have a lot in common with each other and to which we have given the name "giraffe". This is a category in the same sense that the words "mountain" and "snowflake" identify categories of things in nature.

This view of a species as a type, or natural kind, raises the question of whether such things are real. The question is not whether the organisms exist, but whether the *kinds* of organisms exist. There is a school of philosophical thought, called realism that says that natural kinds and other so called *universals* do exist. But what kind of existence would this be? It is one thing to say that a particular giraffe exists, but in what way does the giraffe category exist? This question is the opening for Nominalism which is a philosophical view that types and kinds, and universals in general, do not literally exist.

If the nominalist view is correct then kinds of things, that people have given names to, do not literally exist. It would follow then that because species are named types of organisms, that species do not literally exist. This can be a troubling idea, particularly to a biologist who studies species. If species are not real, then it would not be sensible to talk about "the origin of a species" or the "evolution of a species". As recently at least as the 1950s, some authors adopted this view and wrote of species as not being real.

A useful counterpoint to the nominalist view, in regard to species, was raised by Michael Ghiselin who argued that an individual species is not a type, but rather an actual individual, an actual entity. This idea comes from thinking of a species as an evolving dynamic population. As an entity a species exists quite regardless of whether or not people have observed it and whether or not it has been given a name based on traits shared by the organisms in the species.

Language and the role of human investigators

The nominalist critique of the view that kinds of things exist, raises for consideration the role that humans play in the species problem. For example, Haldane suggested that species are just mental abstractions.

Several authors have noted the similarity between "species", as a word of ambiguous meaning, and points made by Wittgenstein on family resemblance concepts and the indeterminacy of language.

Jody Hey described the species problem as a result of two conflicting motivations by biologists:

1. to categorize and identify organisms;
2. to understand the evolutionary processes that give rise to species.

Under the first view, species appear to us as typical natural kinds, but when biologists turn to understand species evolutionarily they are revealed as changeable and without sharp boundaries. Hey argued that it is unrealistic to expect that one definition of "species" is going to serve the need for categorization and still reflect the changeable realities of evolving species.

Pluralism and monism

Usually it is assumed that biologists approach the species problem with the idea that it would be useful to develop one common viewpoint of species - one single common conception of what species are and of how they should be identified. It is thought that if such a monistic description of species could be developed and agreed upon, then the species problem would be solved.

In contrast some authors have argued for pluralism, claiming that biologists cannot have just one shared concept of species, and that they should accept multiple, seemingly incompatible ideas about species.

David Hull argued that pluralist proposals were unlikely to actually solve the species problem.

Quotations on the species problem

"... I was much struck how entirely vague and arbitrary is the distinction between species and varieties" Darwin 1859 (p. 48)

"No term is more difficult to define than "species," and on no point are zoologists more divided than as to what should be understood by this word". Nicholson (1872) p. 20

"Of late, the futility of attempts to find a universally valid criterion for distinguishing species has come to be fairly generally, if reluctantly, recognized" Dobzhansky (1937) p. 310

"The concept of a species is a concession to our linguistic habits and neurological mechanisms" Haldane (1956)

"The species problem is the long-standing failure of biologists to agree on how we should identify species and how we should define the word 'species'." Hey (2001)

"First, the species problem is not primarily an empirical one, but it is rather fraught with philosophical questions that require - but cannot be settled by - empirical evidence." Pigliucci (2003)

"An important aspect of any species definition whether in neontology or palaeontology is that any statement that particular individuals (or fragmentary specimens) belong to a certain species is an hypothesis (not a fact)"

Chapter- 8

Phylogenetics

In biology, **phylogenetics** is the study of evolutionary relatedness among various groups of organisms (for example, species or populations), which is discovered through molecular sequencing data and morphological data matrices. The term *phylogenetics* is of Greek origin from the terms *phyle/phylon* (φυλή/φῦλον), meaning "tribe, race", and *genetikos* (γενετικός), meaning "relative to birth" from *genesis* (γένεσις, "birth"). Taxonomy, the classification, identification, and naming of organisms, has been richly informed by phylogenetics but remains methodologically and logically distinct.

The fields of phylogenetics and taxonomy overlap in the science of phylogenetic systematics – one methodology of which is called "cladism" or "cladistics" where shared derived characters (synapomorphies) are used to create ancestor-descendant trees (cladograms) and delimit taxa (clades). In biological systematics as a whole, phylogenetic analyses have become essential in researching the evolutionary tree of life.

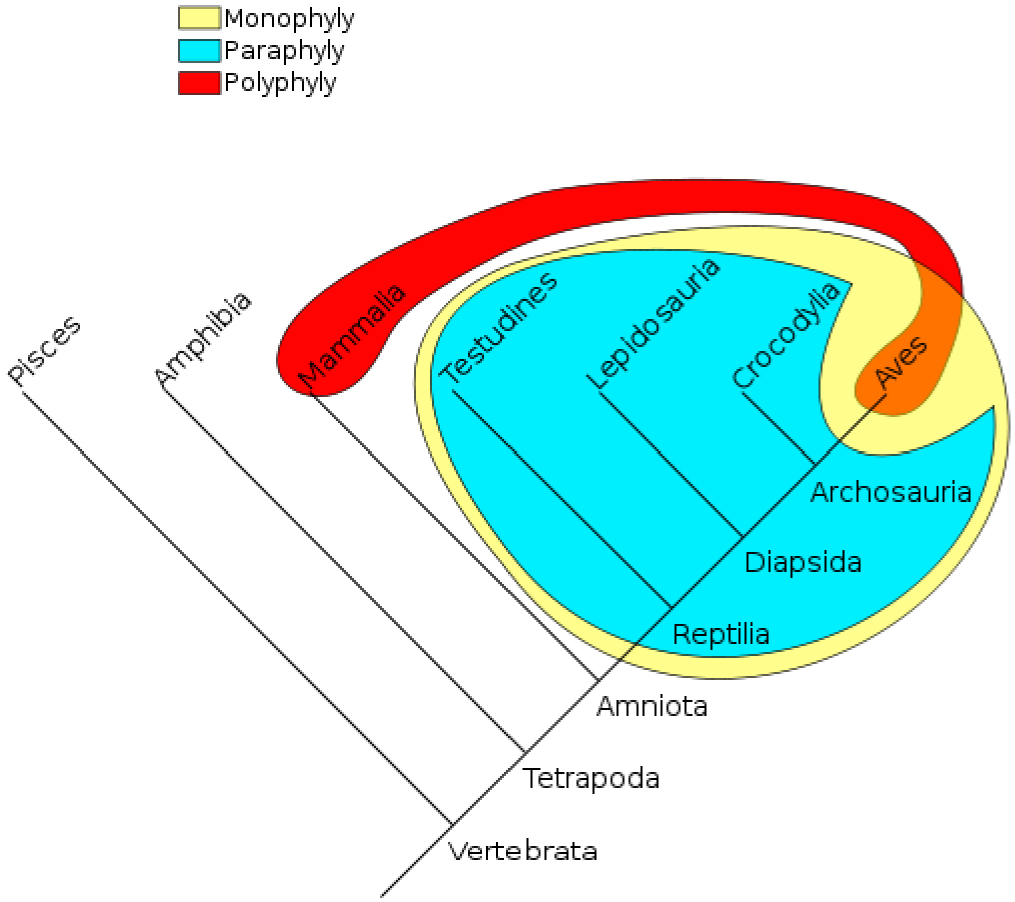
Construction of a phylogenetic tree

Evolution is regarded as a branching process, whereby populations are altered over time and may speciate into separate branches, hybridize together, or terminate by extinction. This may be visualized in a phylogenetic tree.

The problem posed by phylogenetics is that genetic data are only available for living taxa, and the fossil records (osteometric data) contains less data and more-ambiguous morphological characters. A phylogenetic tree represents a hypothesis of the order in which evolutionary events are assumed to have occurred.

Cladistics is the current method of choice to infer phylogenetic trees. The most commonly-used methods to infer phylogenies include parsimony, maximum likelihood, and MCMC-based Bayesian inference. Phenetics, popular in the mid-20th century but now largely obsolete, uses distance matrix-based methods to construct trees based on overall similarity, which is often assumed to approximate phylogenetic relationships. All methods depend upon an implicit or explicit mathematical model describing the evolution of characters observed in the species included, and are usually used for molecular phylogeny, wherein the characters are aligned nucleotide or amino acid sequences.

Grouping of organisms



Phylogenetic groups, or *taxa*, can be monophyletic, paraphyletic, or polyphyletic.

There are some terms that describe the nature of a grouping in such trees. For instance, all birds and reptiles are believed to have descended from a single common ancestor, so this taxonomic grouping (yellow in the diagram below) is called monophyletic. "Modern reptile" (cyan in the diagram) is a grouping that contains a common ancestor, but does not contain all descendants of that ancestor (birds are excluded).

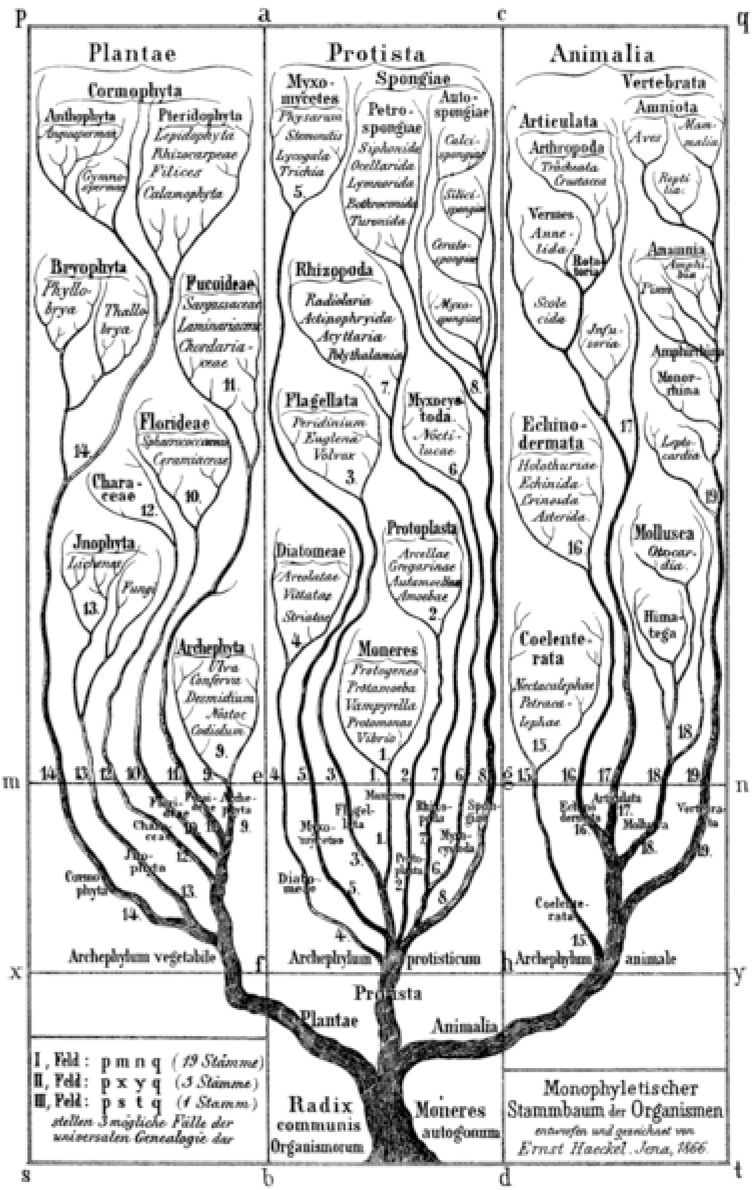
This is an example of a paraphyletic group. A grouping such as warm-blooded animals would include only mammals and birds (red/orange in the diagram) and is called polyphyletic because the members of this grouping do not include the most recent common ancestor.

Molecular phylogenetics

The evolutionary connections between organisms are represented graphically through phylogenetic trees. Due to the fact that evolution takes place over long periods of time that cannot be observed directly, biologists must reconstruct phylogenies by inferring the evolutionary relationships among present-day organisms. Fossils can aid with the reconstruction of phylogenies; however, fossil records are often too poor to be of good help. Therefore, biologists tend to be restricted with analysing present-day organisms to identify their evolutionary relationships. Phylogenetic relationships in the past were reconstructed by looking at phenotypes, often anatomical characteristics. Today, molecular data, which includes protein and DNA sequences, are used to construct phylogenetic trees.

The overall goal of National Science Foundation's Assembling the Tree of Life activity (AToL) is to resolve evolutionary relationships for large groups of organisms throughout the history of life, with the research often involving large teams working across institutions and disciplines. Investigators are typically supported for projects in data acquisition, analysis, algorithm development and dissemination in computational phylogenetics and phyloinformatics. For example, RedToL aims at reconstructing the Red Algal Tree of Life.

Ernst Haeckel's recapitulation theory



Genealogical tree suggested by Haeckel (1866)

During the late 19th century, Ernst Haeckel's recapitulation theory, or biogenetic law, was widely accepted. This theory was often expressed as "ontogeny recapitulates phylogeny", i.e. the development of an organism exactly mirrors the evolutionary development of the species. Haeckel's early version of this hypothesis [that the embryo mirrors *adult* evolutionary ancestors] has since been rejected, and the hypothesis amended as the embryo's development mirroring *embryos* of its evolutionary ancestors. He was accused by five professors of falsifying his images of embryos. Most modern biologists recognize numerous connections between ontogeny and phylogeny, explain them using evolutionary theory, or view them as supporting evidence for that theory. Donald I. Williamson suggested that larvae and embryos represented adults in other taxa that have been transferred by hybridization (the larval transfer theory). However, Williamson's views do not represent mainstream thought in molecular biology, and there is a significant body of evidence against the larval transfer theory.

Gene transfer

In general, organisms can inherit genes in two ways: vertical gene transfer and horizontal gene transfer. Vertical gene transfer is the passage of genes from parent to offspring, and horizontal gene transfer or lateral gene transfer occurs when genes jump between unrelated organisms, a common phenomenon in prokaryotes.

Horizontal gene transfer has complicated the determination of phylogenies of organisms, and inconsistencies in phylogeny have been reported among specific groups of organisms depending on the genes used to construct evolutionary trees.

Carl Woese came up with the three-domain theory of life (eubacteria, archaea and eukaryotes) based on his discovery that the genes encoding ribosomal RNA are ancient and distributed over all lineages of life with little or no horizontal gene transfer. Therefore, rRNAs are commonly recommended as molecular clocks for reconstructing phylogenies.

This has been particularly useful for the phylogeny of microorganisms, to which the species concept does not apply and which are too morphologically simple to be classified based on phenotypic traits.

Taxon sampling and phylogenetic signal

Owing to the development of advanced sequencing techniques in molecular biology, it has become feasible to gather large amounts of data (DNA or amino acid sequences) to infer phylogenetic hypotheses. For example, it is not rare to find studies with character matrices based on whole mitochondrial genomes (~16,000 nucleotides, in many animals). However, it has been proposed that it is more important to increase the number of taxa in the matrix than to increase the number of characters, because the more taxa the more robust is the resulting phylogenetic tree.

This may be partly due to the breaking up of long branches. It has been argued that this is an important reason to incorporate data from fossils into phylogenies where possible. Of course, phylogenetic data that include fossil taxa are generally based on morphology, rather than DNA data. Using simulations, Derrick Zwickl and David Hillis found that increasing taxon sampling in phylogenetic inference has a positive effect on the accuracy of phylogenetic analyses.

Another important factor that affects the accuracy of tree reconstruction is whether the data analyzed actually contain a useful phylogenetic signal, a term that is used generally to denote whether related organisms tend to resemble each other with respect to their genetic material or phenotypic traits. Ultimately, however, there is no way to measure whether a particular phylogenetic hypothesis is accurate or not, unless the "true" relationships among the taxa being examined are already known. The best result an empirical systematist can hope to attain is a tree with branches well-supported by the available evidence.

Importance of missing data

In general, the more data that is available when constructing a tree, the more accurate and reliable the resulting tree will be. Missing data is no less detrimental than simply having less data, although its impact is greatest when most of the missing data is in a small number of taxa. The fewer characters that have missing data, the better; concentrating the missing data across a small number of character states produces a more robust tree.

Role of fossils

Because many morphological characters involve embryological or soft-tissue characters that cannot be fossilized, and the interpretation of fossils is more ambiguous than living taxa, it is sometimes difficult to incorporate fossil data into phylogenies. However, despite these limitations, the inclusion of fossils is invaluable, as they can provide information in sparse areas of trees, breaking up long branches and constraining intermediate character states; thus, fossil taxa contribute as much to tree resolution as modern taxa.

Molecular phylogenies can reveal rates of diversification, but in order to track rates of origination, extinction and patterns in diversification, fossil data must be incorporated. Molecular techniques assume a constant rate of diversification, which is rarely likely to be true; in some (but by no means all) cases, the assumptions inherent in interpreting the fossil record (e.g. a complete and unbiased record) are closer to being true than the assumption of a constant rate, making fossil insights more accurate than molecular reconstructions.

Homoplasy weighting

Certain characters are more likely to be evolved convergently than others; logically, such characters should be given less weight in the reconstruction of a tree. Unfortunately the only objective way to determine convergence is by the construction of a tree – a somewhat circular method. Even so, weighting homoplasious characters does indeed lead to better-supported trees. Further refinement can be brought by weighting changes in one direction higher than changes in another; for instance, the presence of thoracic wings almost guarantees placement among the pterygote insects, although because wings are often lost secondarily, their absence does not exclude a taxon from the group.

Chapter- 9

Cladistics

Cladistics (Ancient Greek: *κλάδος*, *klados*, "branch") is a method of classifying species of organisms into groups called **clades**, which consist of an ancestor organism and all its descendants (and nothing else). For example, birds, dinosaurs, crocodiles, and all descendants (living or extinct) of their most recent common ancestor form a clade. In the terms of biological systematics, a clade is a single "branch" on the "tree of life", a monophyletic group.

Cladistics can be distinguished from other taxonomic systems, such as phenetics, by its focus on shared derived characters (synapomorphies). Systems developed earlier usually employed overall morphological similarity to group species into genera, families and other higher level groups (taxa); cladistic classifications (usually in the form of trees called cladograms) are intended to reflect the relative recency of common ancestry or the sharing of homologous features. Cladistics is also distinguished by an emphasis on parsimony and hypothesis testing (particularly falsificationism), leading to a claim that cladistics is more objective than systems which rely on subjective judgements of relationship based on similarity.

Cladistics originated in the work of the German entomologist Willi Hennig, who referred to it as "phylogenetic systematics" (also the name of his 1966 book); the use of the terms "cladistics" and "clade" was popularized by other researchers. The technique and sometimes the name have been successfully applied in other disciplines: for example, to determine the relationships between the surviving manuscripts of the *Canterbury Tales*.

Cladists use *cladograms* — diagrams which show ancestral relations between species — to represent the monophyletic relationships of species, termed sister-group relationships. This is interpreted as representing phylogeny, or evolutionary relationships. Although traditionally such cladograms were generated largely on the basis of morphological characters, genetic sequencing data and computational phylogenetics are now very commonly used in the generation of cladograms.

Cladistics, either generally or in specific applications, has been criticized from its beginnings. A decision as to whether a particular character is a synapomorphy or not may be challenged as involving subjective judgements, raising the issue of whether cladistics as actually practised is as objective as has been claimed. Formal classifications based on cladistic reasoning are said to emphasize ancestry at the expense of descriptive characteristics, and thus ignore biologically sensible, clearly defined groups which do not fall into clades (e.g. reptiles as traditionally defined or prokaryotes).

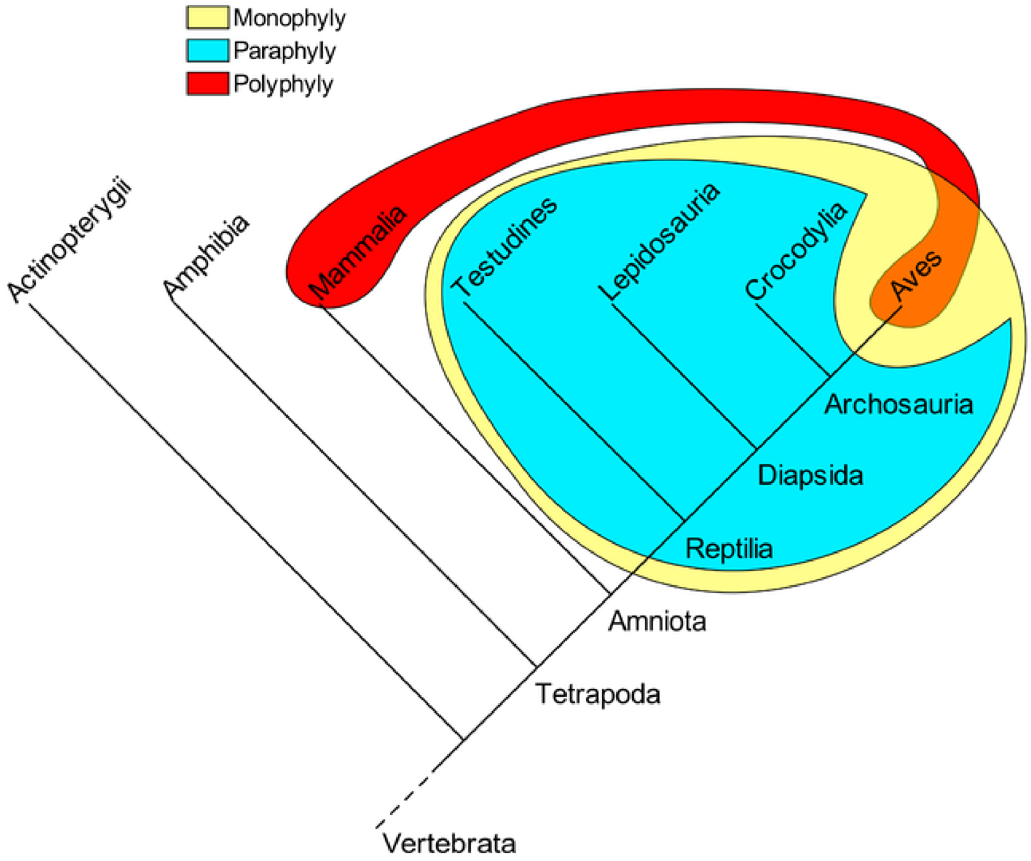
History of cladistics

The term *clade* was introduced in 1958 by Julian Huxley, *cladistic* by Cain and Harrison in 1960, and *cladist* (for an adherent of Hennig's school) by Mayr in 1965. Hennig referred to his own approach as *phylogenetic systematics*. From the time of his original formulation until the end of the 1980s cladistics remained a minority approach to classification. However in the 1990s it rapidly became the dominant method of classification in evolutionary biology. Computers made it possible to process large quantities of data about organisms and their characteristics. At about the same time the development of effective polymerase chain reaction techniques made it possible to apply cladistic methods of analysis to biochemical and molecular genetic features of organisms as well as to anatomical ones.

Cladistics as a successor to phenetics

For some decades in the mid to late twentieth century, a commonly used methodology was phenetics ("numerical taxonomy"). This can be seen as a predecessor to some methods of today's cladistics (namely distance matrix methods such as neighbor-joining), but made no attempt to resolve phylogeny, only similarities.

Clades



Partial Evolutionary Tree of the Vertebrates

The yellow group (sauropsids) is monophyletic, the blue group (traditional reptiles) is paraphyletic, and the red group (warm-blooded animals) is polyphyletic.

A clade is a group of taxa consisting only of an ancestor taxon and all of its descendant taxa. In the diagram provided (a **cladogram**), it is hypothesized that all vertebrates, including ray-finned fishes (Actinopterygii), had a common ancestor, and so form a clade. Within the vertebrates, all tetrapods, including amphibians, mammals, reptiles (as traditionally defined) and birds, are hypothesized to have had a common ancestor, and so also form a clade. The tetrapod ancestor was a descendant of the original vertebrate ancestor, but is not an ancestor of any ray-finned fish living today.

An important caution is that any cladogram is a provisional hypothesis. Although unlikely, future genetic or morphological evidence might suggest that ray-finned fish and amphibians share a common ancestor that was not an ancestor of the other tetrapods. The new information would cause us to define a ray-finned-fish-and-amphibian clade, altering the cladogram.

The relationship between clades can be described in several ways:

- A clade is *basal* to another clade if it contains that other clade as a subset within it. In the example, the vertebrate clade is basal to the tetrapod and ray-finned fish clades. (Some authors have used "basal" differently to mean a clade that is less species-rich than a sister clade, with such a deficit being taken as an indication of 'primitiveness'. Others consider this usage to be incorrect.)
- A clade located within a clade is said to be *nested* within that clade. In the diagram, the tetrapod clade is nested within the vertebrate clade.
- Two clades are *sisters* if they have an immediate common ancestor. In the diagram, crocodiles and birds are sister clades, as are amphibians and amniotes.

Terminology for characters

The following terms are used to identify shared or distinct characters among groups:

- *Plesiomorphy* ("close form") or *ancestral state*, also *symplesiomorphy* ("shared plesiomorphy", i.e. "shared close form"), is a characteristic that is present at the base of a tree (cladogram). Since a plesiomorphy that is inherited from the common ancestor may appear anywhere in a tree, its presence provides no evidence of relationships within the tree. The traditional definition of reptiles (the blue group in the diagram) includes being cold-blooded (i.e. not maintaining a constant high body temperature), whereas birds are warm-blooded. Since cold-bloodedness is a plesiomorphy, inherited from the common ancestor of traditional reptiles and birds, it should not be used to define a group in a system based on cladistics.
- *Apomorphy* ("separate form") or *derived state* is a characteristic believed to have evolved within the tree. It can thus be used to separate one group in the tree from the rest. Within the group which shares the apomorphy it is a *synapomorphy* ("shared apomorphy", i.e. "shared separate form"). For example, within the vertebrates, all tetrapods (and only tetrapods) have four limbs; thus, having four limbs is an synapomorphy for tetrapods. All the tetrapods can legitimately be grouped together because they have four limbs.
- *Homoplasy* is a characteristic shared by members of a tree but not present in their common ancestor. It arises by convergence or reversion. Both mammals and birds are able to maintain a high constant body temperature (i.e. they are 'warm-blooded'). However, the ancestors of each group did not share this character, so it must have evolved independently. Mammals and birds should not be grouped together on the basis that they are warm-blooded.

The terms (sym)plesiomorphy and (syn)apomorphy are relative and their application depends on the position of a group within a tree. An apomorphy of one clade is a plesiomorphy of another contained within it. For example, when trying to decide whether tetrapods should form a clade, an important question is whether having four limbs is a synapomorphy of all the taxa to be included within Tetrapoda: did all the possible members of the Tetrapoda inherit four limbs from a common ancestor, whereas all other vertebrates did not? By contrast, for a group within the tetrapods, such as birds, having four limbs is a plesiomorphy. The fact that ostriches and rheas both have four limbs does not provide any support for putting them into a separate group of 'flightless birds'. Using these two terms allows a greater precision in the discussion of homology, in particular allowing clear expression of the hierarchical relationships among different homologies.

It can be difficult to decide whether a character is in fact the same, and thus can classified as a synapomorphy which may identify a group, or whether it only appears to be the same, and is thus a homoplasy which cannot identify a group. There is a danger of circular reasoning: assumptions about the shape of a phylogenetic tree are used to justify decisions about characters, which are then used as evidence for the shape of the tree. It has been argued that this kind of reasoning has been used by proponents of the view that birds are nested within the theropod dinosaur clade.

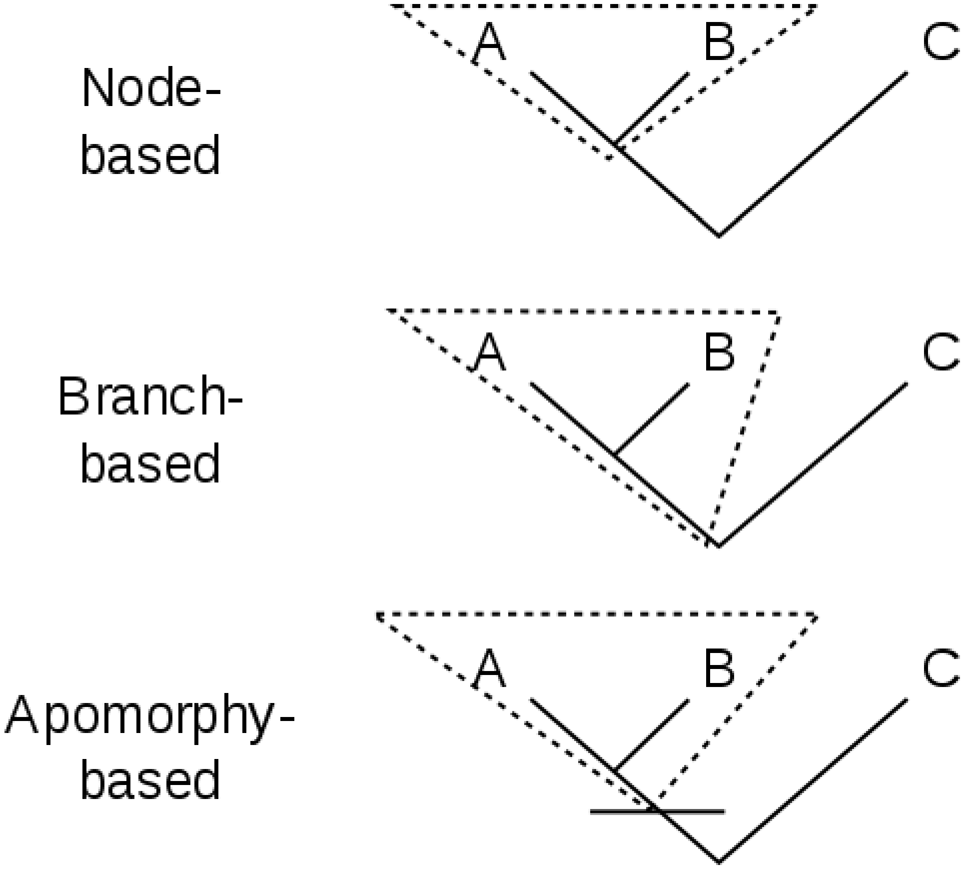
Terminology for groups

Three main types of group can be identified on the basis of their relationships in cladograms. The three can be defined in two different but related ways, as shown in the table below. The first is in terms of the shape of a set of nodes taken from a cladogram. In this approach, an 'ancestor node' is simply a branching point in the diagram; it may or may not correspond to an actual ancestor. The second is in terms of the characters of the taxa being classified and how these characters have been inherited. In this approach, an ancestor is an actual taxon, whether currently known or not.

Term	Node-based definition	Character-based definition
Monophyly	A monophyletic group of nodes in a tree is one which includes all the nodes descended from their most recent common ancestor node, plus the most recent common ancestor node, but no other nodes.	A monophyletic group of taxa is characterized by one or more **synapomorphies**: derived characters inherited by all members of the group from ancestors and not inherited by any other taxa. A monophyletic group is a 'clade'. A 'crown group' is an example of a monophyletic group.
Paraphyly	A paraphyletic group of nodes in a tree is one which is constructed by taking a monophyletic group and removing one or more smaller monophyletic groups. (Removing one group produces a singly paraphyletic group, removing two a doubly paraphylectic group, and so on.) A paraphyletic group is necessarily non-monophyletic.	A paraphyletic group of taxa is characterized by one or more **(sym)plesiomorphies**: characters inherited from ancestors but not present in all of their descendants. As a consequence, a paraphyletic group is truncated, in that it excludes one or more monophyletic taxa from an initially monophyletic group. An alternative name is an 'evolutionary grade', refering to the ancestral character state within the group. A 'stem group' is an example of a

		paraphyletic group.
Polyphyly	A polyphyletic group of nodes in a tree is one which is neither monophyletic nor paraphyletic.	A polyphyletic group of taxa is characterized by by one or more **homoplasies**: characters which have converged or reverted so as to appear to be the same but which have not been inherited from common ancestors. As a consequence, polyphyletic groups of taxa are totally artificial.

Branch-based definitions of clade



Three alternative ways to define a clade

The node-based definition of a monophyletic group (i.e. a clade) given above regards the lines in the cladogram only as a way of showing connections between taxa. This is appropriate when considering only living (extant) taxa; however, when extinct taxa are to be included in a cladogram, lines correspond to sequences of ancestors. There are two alternative ways of defining a clade which explicitly take into account the line below the branching point at the base of a clade. These definitions are most notably set out in the PhyloCode.

Consider how a clade combining A and B in the diagram can be defined.

- *Node-based*: The node-based definition specifies A+B as the *last* common ancestor of A and B, and all descendants of that ancestor. It thus excludes from the clade the line below the junction of A and B. Crown groups are a type of node-based clade.
- *Branch-based*: A branch-based definition specifies A+B as the *first* ancestor of A which is not also an ancestor of C, and all descendants of that ancestor. It thus includes in the clade the line below the junction of A and B. (This type of definition was originally called "stem-based", but this was changed to avoid confusion with the term "stem group", which is parapyletic.) Total groups are a type of branch-based clade.
- *Apomorphy-based*: An apomorphy-based definition specifies A+B as the first ancestor of A to possess derived trait M homologously (that is, synapomorphically) with that trait in A, and all descendants of that ancestor. It thus includes in the clade only that part of the line below the junction of A and B which corresponds to ancestors possessing the apomorphy. The process of identifying and naming groups based on apomorphies is the method that most resembles classical systematics, with the proviso that cladistic taxa always denote a clade.

Note that these alternative definitions do not alter the classification of the tips of the tree, and so are equivalent if only living (extant) taxa are being considered.

Cladograms

Cladists use *cladograms*, diagrams which show ancestral relations between taxa, to represent the evolutionary tree of life. Although traditionally such cladograms were generated largely on the basis of morphological characters, molecular sequencing data and computational phylogenetics are now very commonly used in the generation of cladograms.

The starting point of cladistic analysis is a group of species and molecular, morphological, or other data characterizing those species. The end result is a tree-like relationship diagram called a cladogram, or sometimes a *dendrogram* (Greek for "tree drawing"). The cladogram graphically represents a hypothetical evolutionary process. Cladograms are subject to revision as additional data become available.

The terms "evolutionary tree", and sometimes "phylogenetic tree" are often used synonymously with cladogram but others treat phylogenetic tree as a broader term that includes trees generated with a nonevolutionary emphasis. In cladograms, all species lie at the leaves. The two taxa on either side of a split, with a common ancestor and no additional descendents, are called "sister taxa" or "sister groups". Each subtree, whether it contains only two or a hundred thousand items, is called a "clade". Many cladists require that all forks in a cladogram be 2-way forks. Some cladograms include 3-way or 4-way forks when there are insufficient data to resolve the forking to a higher level of detail.

For a given set of taxa, the number of distinct cladograms that can be drawn (ignoring which cladogram best matches the taxon characteristics) is:

Number of taxa	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**N**
Number of rooted cladograms	1	3	15	105	945	10,395	135,135	2,027,025	34,459,425	1*3*5*7*...*(2N-3)

This superexponential growth of the number of possible cladograms explains why manual creation of cladograms becomes very difficult when the number of taxa is large. If a cladogram represents N taxa, the number of levels (the "depth") in the cladogram is on the order of $\log_2(N)$. For example, if there are 32 species of deer, a cladogram representing deer could be around 5 levels deep (because $2^5 = 32$), although this is really just the lower limit. A cladogram representing the complete tree of life, with about 10 million species, could be about 23 levels deep. This formula gives a lower limit, with the actual depth generally a larger value, because the various branches of the cladogram will not be uniformly deep. Conversely, the depth may be shallower if forks larger than 2-way forks are permitted.

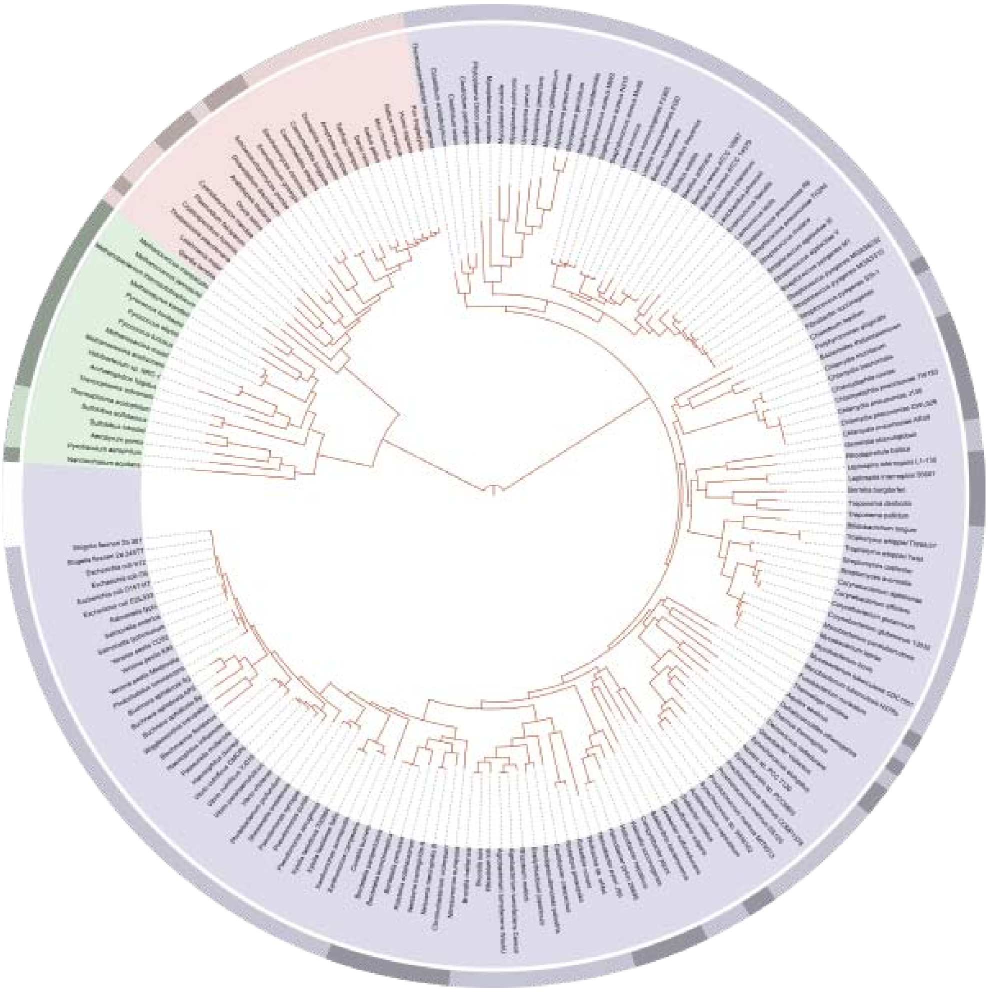
A cladogram tree has an implicit time axis, with time running forward from the base of the tree to the leaves of the tree. If the approximate date (for example, expressed as millions of years ago) of all the evolutionary forks were known, those dates could be captured in the cladogram. Thus, the time axis of the cladogram could be assigned a time scale (e.g. 1 cm = 1 million years), and the forks of the tree could be graphically located along the time axis. Such cladograms are called *scaled cladograms*. Many cladograms are not of this type, for a variety of reasons:

- They are built from species characteristics that cannot be readily dated (e.g. morphological data in the absence of fossils or other dating information)
- When the characteristic data are DNA/RNA sequences, it is feasible to use sequence differences to establish the relative ages of the forks, but converting those ages into actual years requires a significant approximation of the rate of change
- Even when the dating information is available, positioning the cladogram's forks along the time axis in proportion to their dates may cause the cladogram to become difficult to understand or hard to fit within a human-readable format

Cladistics makes no distinction between extinct and extant species, and it is appropriate to include extinct species in the group of organisms being analyzed. Cladograms that are based on DNA/RNA generally do not include extinct species because DNA/RNA samples from extinct species are rare. Cladograms based on morphology, especially morphological characteristics that are preserved in fossils, are more likely to include extinct species.

Cladistics in taxonomy

Phylogenetic nomenclature contrasted with traditional taxonomy



A highly resolved, automatically generated tree of life based on completely sequenced genomes

Most taxonomists have used the traditional approaches of Linnaean taxonomy and later Evolutionary taxonomy to organize life forms. These approaches use several fixed levels of a hierarchy, such as kingdom, phylum, class, order, and family. Phylogenetic nomenclature does not feature those terms, because the evolutionary tree is so deep and so complex that it is inadvisable to set a fixed number of levels.

Evolutionary taxonomy insists that groups reflect phylogenies. In contrast, Linnaean taxonomy allows both monophyletic and paraphyletic groups as taxa. Since the early 20th century, Linnaean taxonomists have generally attempted to make at least family- and lower-level taxa (i.e. those regulated by the codes of nomenclature) monophyletic. Ernst Mayr in 1985 drew a distinction between the terms cladistics and phylogeny: "It would seem to me to be quite evident that the two concepts of phylogeny (and their role in the construction of classifications) are sufficiently different to require terminological distinction. The term *phylogeny* should be retained for the broad concept of phylogeny, promoted by Darwin and adopted by most students of phylogeny in the ensuing 90 years. The concept of phylogeny as mere genealogy should be terminologically distinguished as *cladistics*. To lump the two concepts together terminologically could not help but produce harmful equivocation."

Willi Hennig's pioneering work provoked a spirited debate about the relative merits of phylogenetic nomenclature versus Linnaean or evolutionary taxonomy, which has continued down to the present; however Hennig did not advocate abandoning the Linnaean nomenclatural system. Some of the debates in which the cladists were engaged had been running since the 19th century, but they were renewed fervor, as can be seen from the *Foreword* to Hennig (1979) by Rosen, Nelson, and Patterson:

"Encumbered with vague and slippery ideas about adaptation, fitness, biological species and natural selection, neo-Darwinism (summed up in the "evolutionary" systematics of Mayr and Simpson) not only lacked a definable investigatory method, but came to depend, both for evolutionary interpretation and classification, on consensus or authority."

Phylogenetic nomenclature strictly and exclusively follows phylogeny and has arbitrarily deep trees with binary branching: each taxon corresponds to a clade. Linnaean taxonomy, while since the advent of evolutionary theory following phylogeny, also may subjectively consider similarity and has a fixed hierarchy of taxonomic ranks, and its taxa are not required to correspond to clades.

Paraphyletic groups discouraged

Many cladists discourage the use of paraphyletic groups in classification of organisms, because they detract from cladistics' emphasis on clades (monophyletic groups). In contrast, proponents of the use of paraphyletic groups argue that any dividing line in a cladogram creates both a monophyletic section above and a paraphyletic section below. They also contend that paraphyletic taxa are necessary for classifying earlier sections of the tree – for instance, the early vertebrates that would someday evolve into the family Hominidae cannot be placed in any other monophyletic family. They also argue that paraphyletic taxa provide information about significant changes in organisms' morphology, ecology, or life history – in short, that both paraphyletic groups and clades are valuable notions with separate purposes.

Complexity of the Tree of Life

The cladistic tree of life is a fractal:

"The tree of life is inherently fractal-like in its complexity, ... Look closely at the 'lineage' of a phylogeny ... and it dissolves into many smaller lineages, and so on, down to a very fine scale."

The overall shape of a dichotomous (bifurcating) tree is recursive; as a viewpoint zooms into the tree of life, the same type of tree appears no matter what the scale. When extinct species are considered (both known and unknown), the complexity and depth of the tree can be very large. Moreover the tree continues to recreate itself by bifurcation, a series of events called fractal evolution. Every single speciation event, including all the species that are now extinct, represents an additional fork on the hypothetical, complete cladogram of the tree of life.

The tree of life is a quasi-self-similar fractal; that is, the deep reconstruction is not as regular as the shallow reconstruction. By shallow Mishler means the most recent branching toward and at the tips, and by deep the more ancient branches further back, which are harder to reconstruct and are missing unknown extinct lines. In the shallow part of the tree, branching events are relatively regular; it is often possible to estimate the times between them. In the deep part of the tree, "homology assessments" are "difficult" and the times vary widely. At this level Eldredge's and Gould's punctuated equilibrium applies, which hypothesizes long periods of stability followed by punctuations of rapid speciation, based on the fossil record.

PhyloCode approach to naming species

A formal code of phylogenetic nomenclature, the PhyloCode, is currently under development. It is intended for use by both those who would like to abandon Linnaean taxonomy and those who would like to use taxa and clades side by side. In several instances it has been employed to clarify uncertainties in Linnaean systematics so that in combination they yield a taxonomy that unambiguously places problematic groups in the evolutionary tree in a way that is consistent with current knowledge.

Example

For example, Linnaean taxonomy contains the taxon Tetrapoda, defined morphologically as vertebrates with four limbs (as well as animals with four-limbed ancestors, such as snakes), which is often given the rank of superclass, and divides into the classes Amphibia, Reptilia, Aves, Mammalia.

Phylogenetic nomenclature also contains the taxon Tetrapoda, whose living members can be classified phylogenically as "the clade defined by the common ancestor of amphibians and mammals", or more precisely the clade defined by the common ancestor of a specific amphibian and mammal (or bird or snake). This definition gives us the Crown group tetrapods (or Crown-Tetrapoda). A few primitive four legged ancestors (the Ichthyostegalia) fall outside Crown-Tetrapoda. An alternative is to define tetrapoda as all animals more closely related to mammals than to lungfish (our nearest living non-tetrapod relatives). In this definition, the ichthyostegalians are included, together with a host of fossil animals usually classed as crossopterygian fish. This wider definition is termed Pan-Tetrapoda. A third option is to define Tetrapoda according to their apomorphy (their unique trait, i.e. having legs rather than fins), a definition that yield the same group as the Linnaean taxon.

Non of the phylogenetic taxa as described above have a rank, and neither do its subtaxa. All the subclades are contained within one another. The clades are not divided into several non-overlapping taxa (as in traditional taxonomy), rather the clade is split into two clades at the first branching, a process repeated throughout. With regards to the traditional classes, Aves and Mammalia are subclades, contained in the subclade Amniota, while Reptilia and Amphibia are paraphyletic taxa, not clades. Instead of classifying non-mammalian, non-avian amniotes as reptiles, Amniota is divided into the two clades Sauropsida (which contains birds and all living amniotes other than mammals, including all living traditional reptiles) and Theropsida (mammals and the extinct mammal-like reptiles). Similarly, Amphibia can be split into the Batrachomorpha (fossil amphibians more closely related to modern amphibians) and Reptiliomorpha, the latter of which the amiotes is a sub-clade. Ichthyostegalians and other Stem-tetrapods represent sister groups from splits predating the Batrachomorpha/Reptilopmorpha split.

Summary of advantages of phylogenetic nomenclature

Proponents of phylogenetic nomenclature enumerate key distinctions between phylogenetic nomenclature and Linnaean taxonomy as follows:

Phylogenetic Nomenclature	**Linnaean Taxonomy**
Handles arbitrarily deep trees.	Often must invent new level names (such as superorder, suborder, infraorder, parvorder, magnorder) to accommodate new discoveries. Biased towards trees about 4 to 12 levels deep.
Discourages naming or use of groups that are not monophyletic	Acceptable to name and use paraphyletic groups
Primary goal is to reflect actual process of evolution	Primary goal is to group species based on morphological similarities
Assumes that the shape of the tree will change frequently with new discoveries	New discoveries often require renaming or releveling of Classes, Orders, and Kingdoms

Summary of criticisms of phylogenetic nomenclature

Critics of phylogenetic nomenclature include Ashlock, Mayr, and Williams. Some of their criticisms include:

Phylogenetic Nomenclature	**Linnaean Taxonomy**
Limited to entities related by evolution or ancestry	Supports groupings without reference to evolution or ancestry
Does not include a process for naming	Includes a process for giving unique names

species	to species
Clade definitions emphasize ancestry at the expense of descriptive characteristics	Taxa definitions based on tangible characteristics
Ignores sensible, clearly defined paraphyletic groups such as reptiles	Permits clearly defined groups such as reptiles
Difficult to determine if a given species is in a clade or not (e.g. if clade X is defined as "most recent common ancestor of A and B along with its descendants", then the only way to determine if species Y is in the clade is to perform a complex evolutionary analysis)	Straightforward process to determine if a given species is in a taxon or not
Limited to organisms that evolved by inherited traits; not applicable to organisms that evolved via complex gene sharing or lateral transfer	Applicable to all organisms, regardless of evolutionary mechanism

Application to other disciplines

The comparisons used to acquire data on which cladograms can be based are not limited to the field of biology. Any group of individuals or classes, hypothesized to have a common ancestor, and to which a set of common characteristics may or may not apply, can be compared pairwise. Cladograms can be used to depict the hypothetical descent relationships within groups of items in many different academic realms. The only requirement is that the items have characteristics that can be identified and measured.

Recent attempts to use cladistic methods outside of biology address the reconstruction of lineages in:

- Anthropology and archeology. Compares cultures or artifacts using groups of cultural traits or artifact features.
- Linguistics. Compares languages using groups of linguistic features.
- Textual criticism or Stemmatics. Compares manuscripts of the same work (original lost) using groups of distinctive copying errors.
- Ethology. Compares animal species using behavioral traits presumed hereditary.

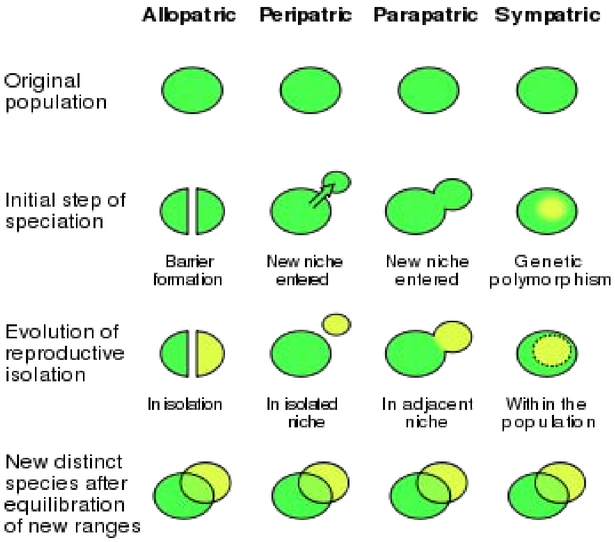
Chapter- 10

Allopatric Speciation

Allopatric speciation (from the ancient Greek allos, "other" + Greek patrida, "fatherland") or **geographic speciation** is speciation that occurs when biological populations of the same species become isolated due to geographical changes such as mountain building or social changes such as emigration. The isolated populations then undergo genotypic and/or phenotypic divergence as: (a) they become subjected to different selective pressures, (b) they independently undergo genetic drift, and (c) different mutations arise in the populations' gene pools.

The separate populations over time may evolve distinctly different characteristics. If the geographical barriers are later removed, members of the two populations may be unable to successfully mate with each other, at which point, the genetically isolated groups have emerged as different species. Allopatric isolation is a key factor in speciation and a common process by which new species arise. Adaptive radiation, as observed by Charles Darwin in Galapagos finches, is a consequence of allopatric speciation among island populations.

Isolating Mechanisms



Comparison of **allopatric**, peripatric, parapatric and sympatric speciation

Allopatric speciation may occur when a species is subdivided into two genetically isolated populations. *Allopatric* and *allopatry* are terms from biogeography, referring to organisms whose ranges are entirely separate such that they do not occur in any one place together. If these organisms are closely related (e.g. sister species), such a distribution is usually the result of *allopatric speciation*. Separation may be attributed to either geological processes or population dispersal.

Geographical isolation

Geological processes can fragment a population through such events as emergence of mountain ranges, canyon formation, glacial processes, the formation or destruction of land bridges, or the subsidence of large bodies of water. On a global scale, plate tectonics is a major geological factor leading to separation of populations and the resulting distribution of species.

Approximately 50,000 years ago, the Death Valley region of the western United States had a rainy climate which produced an interconnecting system of freshwater rivers and lakes. Climatic changes resulted in a drying trend that has continued for the last 10,000 years. As the lakes and rivers shrunk, fish populations became geologically isolated. The few remaining (separated) springs are currently home to a variety of fish, many sharing a close common ancestor; yet each has uniquely adapted to its own particular pool.

The extent to which a geological barrier can effectively isolate a population correlates to the mobility of the organism or its offspring. For example physical barriers such as canyons may effectively block migration and dispersal of small mammals; however, have little impact on flying birds or wind-borne seeds.

Population Dispersal

Population dispersal is used to describe migratory events, either in the form of range expansion (natural movement away from parents) or jump dispersal (crossing of barriers), which may lead to genetic isolation. If the smaller population fragment becomes genetically isolated from the parental group, it may be subjected to its own unique mutations, selection forces, and genetic drift effects; thus, it will follow its own evolutionary pathway. Migrations or accidental relocations (such as birds being blown off course) may lead to population fragments; whereby groups merely become separated by distance. Once gene flow between the two groups is disrupted, speciation becomes a possibility.

Allopatric speciation in peripheral populations

When populations become genetically isolated, heritable variations may accumulate so that they become different from the parental population. Given sufficient time, these variations may lead to reproductive isolation.

Portions of a populations that exist along the edges of the parent population's geographic territory have higher likelihood of developing reproductive isolation. Such peripheral populations are likely to possess genes that are different from the parental population. After isolation, the founding population is less likely to represent the gene pool of the parent population. In addition, peripheral isolates are likely to represent a small number of individuals, meaning their gene pool is more susceptible to the effects of genetic drift (random chance). Furthermore, it is likely that the peripheral population will inhabit an environment different from its ancestral gene pool, likely causing it to be subjected to different selective pressures as it colonizes new areas. The outer periphery of a population's habitat tends to be extreme; hence, the reason range expansion is kept in check. For most peripheral isolates, it is more likely that they die off rather than survive and speciate.

Genesis of reproductive barriers

Adaptive divergence may occur when a population becomes geographically divided: followed by an accumulation of genetic differences as they adapt to their own unique environments. Reproductive barriers do not evolve as a consequence of external forces that drive populations toward speciation. Rather, the evolution of reproductive isolation, leading to speciation, is generally thought to be an incidental by-product of genetic divergence, particularly adaptive changes that evolve through natural selection in response to different environmental conditions in separate geographic areas.

The Biological Species Concept, proposed by Ernst Mayr, in the 1942, emphasizes reproductive isolation as the basis of defining a species. The definition states: "A species is defined as a population or group of populations whose members have the potential to interbreed with one another in nature and to produce viable offspring, but cannot produce viable, fertile offspring with members of other species." Mayr, a proponent of allopatric speciation, hypothesized that adaptive genetic changes that accumulate between allopatric populations cause negative epistasis in hybrids, resulting in sterility of the offspring.

If there is considerable genetic and phenotypic change without the loss of the capacity for interbreeding, then such hybridization is simply prevented by the geographical separation of populations. In this case the populations are normally regarded as subspecies.

The frequency of other types of speciation, such as sympatric speciation, parapatric speciation, and heteropatric speciation, is debated. Proponents of peripatric speciation contend that small population size in the peripheral isolate (sometimes referred to as a "splinter population") increases genetic drift, which can be a more powerful force than natural selection in small populations. It deconstructs complex genotypes, allowing the creation of novel gene combinations. Both forms need not be mutually exclusive. In practice, passive isolation or fragmentation as well as active dispersal seem to play a role in many cases of speciation.

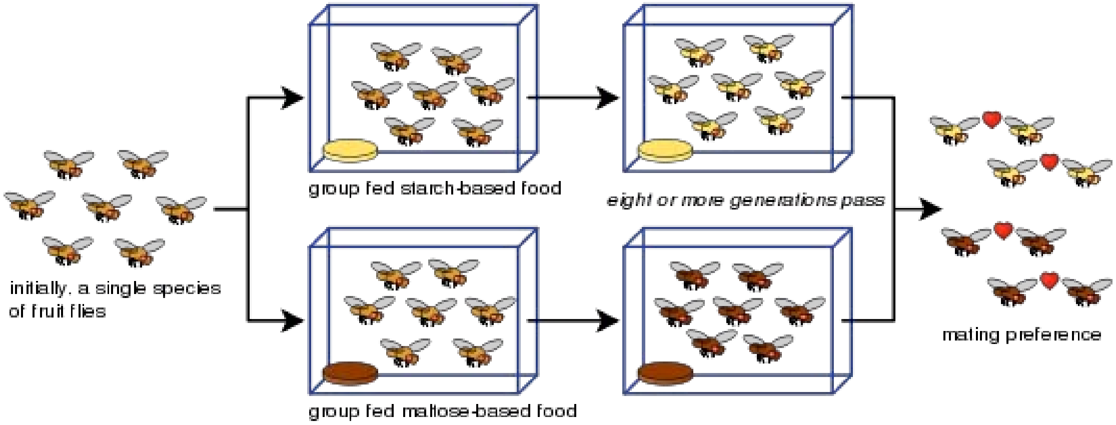
Alternative Modes of Speciation

Sympatric speciation represents an alternative method of speciation that does not require physical separation; instead speciation occurs within a population sharing the same geographic boundaries. For example, the development of polyploidy in plant species can lead to a new species arising within the geographic range of its parent population.

In parapatric speciation there is no physical barrier to gene exchange within the population. Instead, the population is continuous; however, mating is not random. Individuals mate with their closest neighbors rather than with individuals in a more distant location. Divergence may occur as a consequence of both reduced gene flow and natural selection, imposed by the large distance between individuals within a population's habitat.

Allopatric speciation is thought to be the dominant mode of speciation.

Examples



Allopatric speciation among fruit flies

The African Elephant has always been regarded as a single species but, because of morphological and DNA differences, some scientists classify them into three subspecies. Researchers at the University of California, San Diego have argued that divergence due to geographical isolation has gone further, and the elephants of West Africa should be regarded as a separate species from both the savanna elephants of Central, Eastern and Southern Africa, or the forest elephants of Central Africa. A similar situation exists with the Asian Elephant, which has four distinct living sub-species.

Other cases arise where two populations that are quite distinct morphologically, and are native to different continents, have been classified as different species; but when members of one species are introduced into the other's range, they are found to interbreed freely, showing that they were in fact only geographically isolated subspecies. This was found to be the case when the Mallard was introduced into New Zealand and interbred freely with the native Grey Duck, which had been classified as a separate species. It is controversial whether its specific status can now be retained.

Chapter- 11

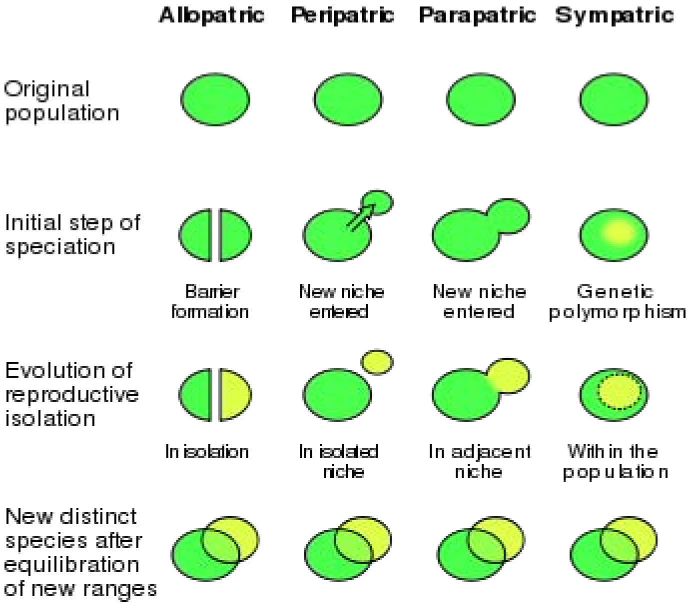
Sympatric Speciation

A speciation in which new species evolve from a single ancestral species while inhabiting the same geographic region. In evolutionary biology and biogeography, **sympatric** and **sympatry** are terms referring to organisms whose ranges overlap or are even identical, so that they occur together at least in some places. If these organisms are closely related (e.g. sister species), such a distribution may be the result of **sympatric speciation**. Etymologically, sympatry is derived from the Greek roots *συν* (together, with) and *πατρίς* (homeland or fatherland). The term was invented by Poulton in 1904, who explains the derivation.

Sympatry is one of four theoretical models for the phenomenon of speciation. In contrast to allopatry, populations undergoing sympatric speciation are not geographically isolated by, for example, a mountain or a river.

In multicellular eukaryotic organisms, sympatric speciation is thought to be an uncommon but plausible process by which genetic divergence (through reproductive isolation) of various populations from a single parent species and inhabiting the same geographic region leads to the creation of new species. In bacteria, however, the analogous process (defined as "the origin of new bacterial species that occupy definable ecological niches") is more common and occurs through horizontal gene transfer.

Evidence



Comparison of allopatric, peripatric, parapatric and sympatric speciation.

A number of models have been proposed to account for this mode of speciation. The most popular, which invokes the disruptive selection model, was first put forward by John Maynard Smith in 1966. Maynard Smith suggested that homozygous individuals may, under particular environmental conditions, have a greater fitness than those with alleles heterozygous for a certain trait. Under the mechanism of natural selection, therefore, homozygosity would be favoured over heterozygosity, eventually leading to speciation. Sympatric divergence could also result from the sexual conflict.

Disruption may also occur in multiple-gene traits. The Medium Ground Finch (*Geospiza fortis*) is showing gene pool divergence in a population on Santa Cruz Island. Beak morphology conforms to two different size ideals, while intermediate individuals are selected against. Some characteristics (termed magic traits) such as beak morphology may drive speciation because they also affect mating signals. In this case, different beak phenotypes may result in different bird calls, providing a barrier to exchange between the gene pools.

A well studied circumstance of sympatric speciation is when insects feed on more than one species of host plant. In this case insects become specialized as they struggle to overcome the various plants' defense mechanisms. (Drès and Mallet, 2002)

Rhagoletis pomonella, the apple maggot, may be currently undergoing sympatric or, more precisely, heteropatric speciation. The apple feeding race of this species appears to have spontaneously emerged from the hawthorn feeding race in the 1800 - 1850 AD time frame, after apples were first introduced into North America. The apple feeding race does not now normally feed on hawthorns, and the hawthorn feeding race does not now normally feed on apples. This may be an early step towards the emergence of a new species. Isolated and relatively homogeneous habitats such as crater lakes and islands are among the best geographical settings in which to demonstrate sympatric speciation. For example, Nicaragua crater lake cichlid fishes include at least one species that has evolved by sympatric speciation

Allochrony offers some empirical evidence that sympatric speciation has taken place, as many examples exist of recently diverged (sister taxa) allochronic species.

Sympatric speciation events are vastly more common in plants, as they are prone to developing multiple homologous sets of chromosomes, resulting in a condition called polyploidy. The polyploidal offspring occupy the same environment as the parent plants (hence sympatry), but are reproductively isolated.

A rare example of sympatric speciation in animals is the divergence of "resident" and "transient" Orca forms in the northeast Pacific. Resident and transient orcas inhabit the same waters, but avoid each other and do not interbreed. The two forms hunt different prey species and have different diets, vocal behaviour, and social structures. Some divergences between species could also result from contrasts in microhabitats.

The polecat *Mustela putorius* exhibited a rare dark phenotype similar to the European mink *Mustela lutreola* phenotype which is directely influenced by peculiarities of forest brooks.

Controversy

Debated almost since the beginning of popular evolutionary thought, sympatric speciation is still a highly contentious issue. By 1980 the theory was largely unfavourable given the void of empirical evidence available, and more critically the conditions scientists expect to be required. Ernst Mayr, one of the foremost thinkers on evolution, completely rejected sympatry outright, ushering in a climate of hostility towards the theory. While still debatable, well documented empirical evidence now exists, and the development of sophisticated theories incorporating multilocus genetics have followed.