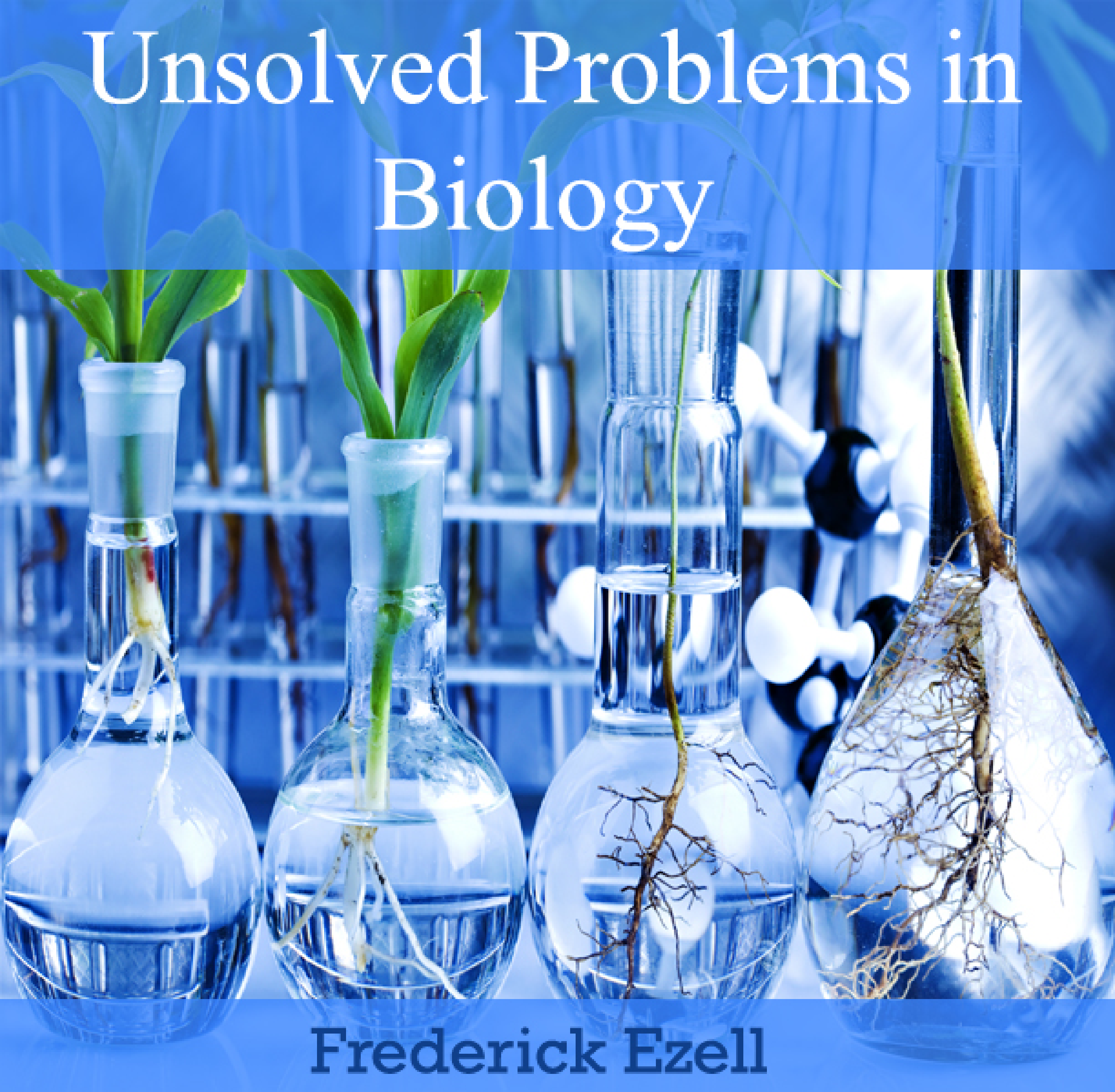
Unsolved Problems in Biology



Frederick Ezell

First Edition, 2012

ISBN 978-81-323-4250-2

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

Table of Contents

Chapter 1

Arthropod Head Problem

The **arthropod head problem** is a long-standing zoological dispute concerning the segmental composition of the heads of the various arthropod groups, and how they are evolutionarily related to each other. While the dispute has historically centered on the exact make-up of the insect head, it has been widened to include other living arthropods such as the crustaceans and chelicerates; and fossil forms, such as the many arthropods known from exceptionally-preserved Cambrian faunas. While the topic has classically been based on insect embryology, in recent years a great deal of developmental molecular data have become available. Dozens of more or less distinct solutions to the problem, dating back to at least 1897, have been published, including several in the 2000s.

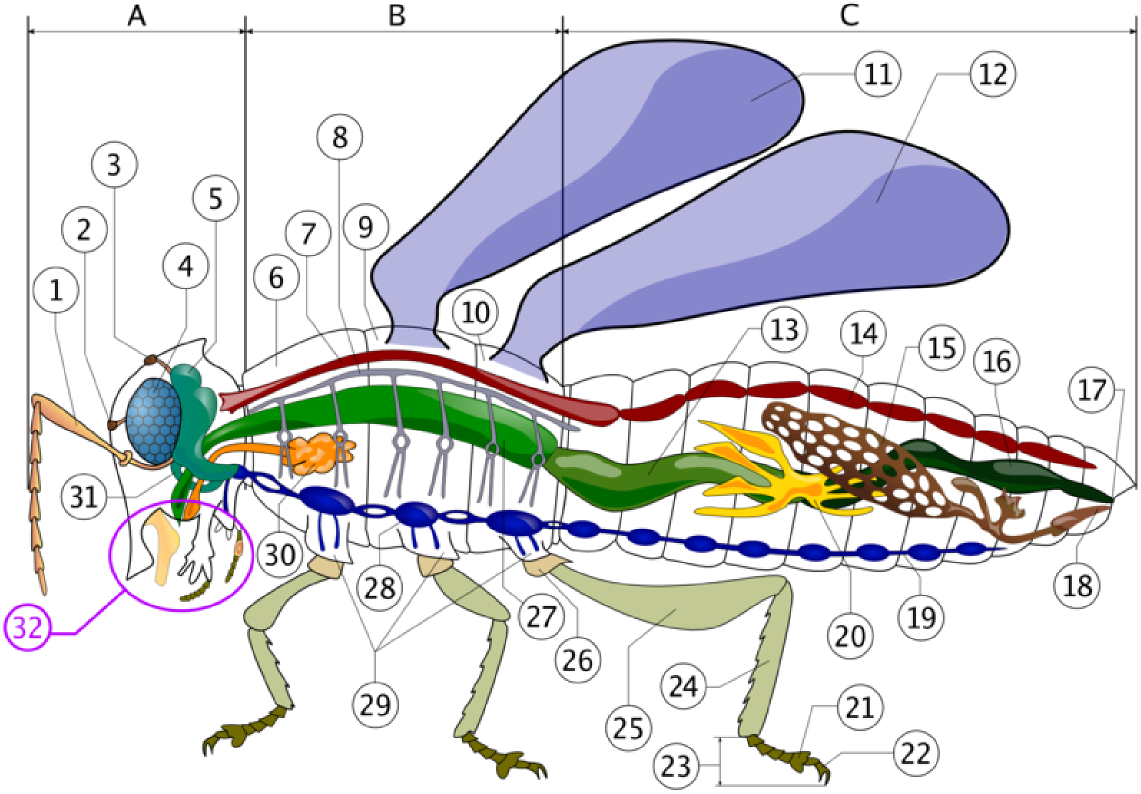
The arthropod head problem is popularly known as the "endless dispute", the title of a famous paper on the subject by Jacob G. Rempel in 1975, referring to its apparently intractable nature. Although some progress has been made since that time, the precise nature of especially the labrum and the pre-oral region of arthropods remain highly controversial.

Background

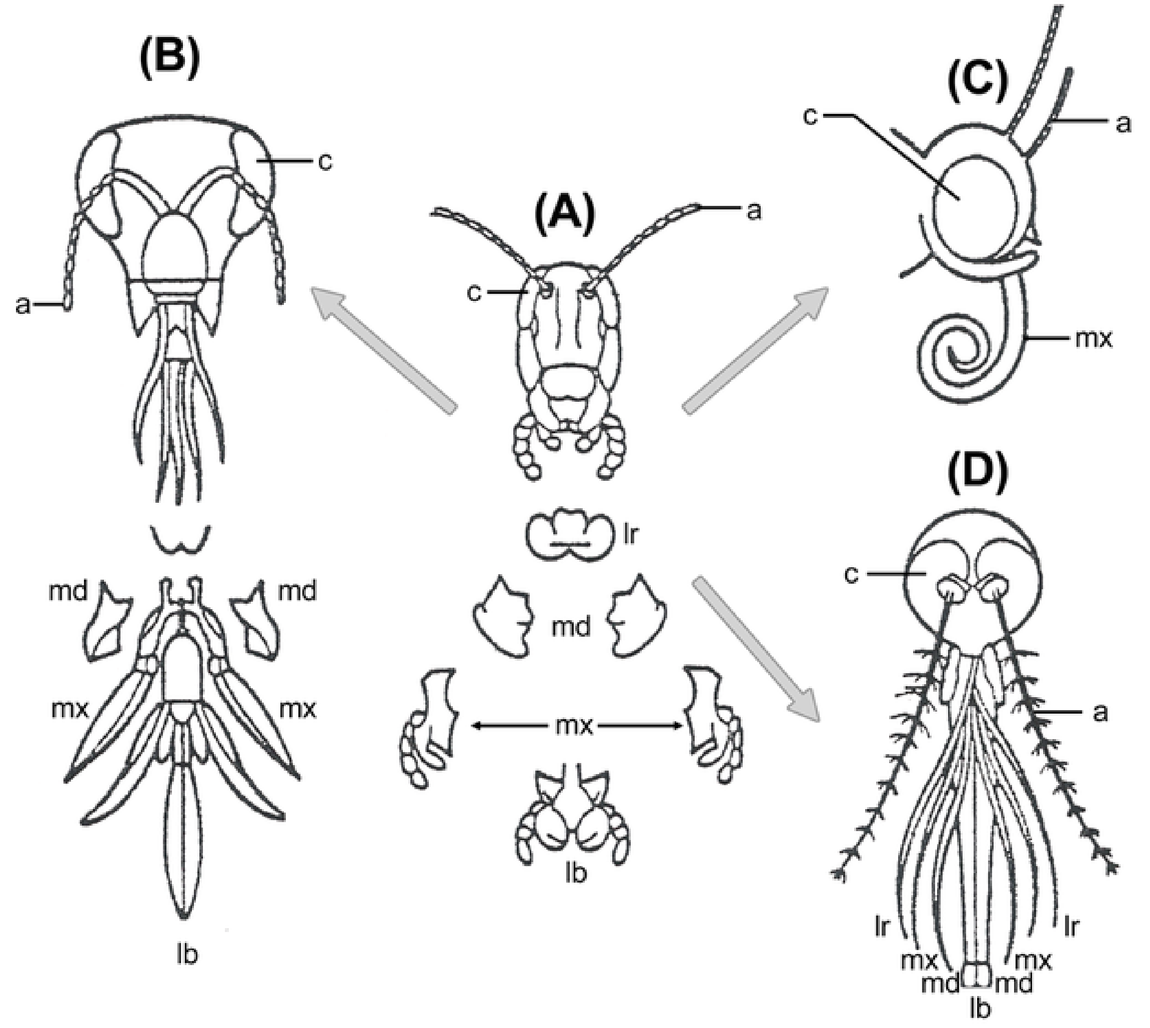
> It would be too bad if the question of head segmentation ever should be finally settled; it has been for so long such fertile ground for theorizing that arthropodists would miss it as a field for mental exercise.
> R.E. Snodgrass, 1960

According to recent research, it has been suggested that some key events in the evolution of the arthropod body resulted from changes in certain Hox genes' DNA sequences. The trunks of arthropods are built out of repeated segments, which are typically associated with various structures such as a pair of appendages, apodemes for muscle attachment, ganglia and, at least embryologically, coelomic cavities. While many arthropod segments are modified to a greater or lesser extent (for example, only three of the insect thorax and abdominal segments typically bear appendages), it is widely assumed that the ancestral state was for all of the segments to be nearly identical. However, while the segmental organisation of the trunks of adult arthropods can be usually readily seen, that of the head is much less obvious. Arthropod heads are typically fused capsules that bear a variety of complex structures such as the eyes, antennae and mouth parts. The challenge that the arthropod head problem has to address is to what extent the various structures of the arthropod head can be resolved into a set of hypothetical ancestral segments. Given the high compaction and complexity of adult arthropod heads, much attention has been directed towards understanding the developmental processes that give rise to them, in the hope that they will reveal their segmental organisation more clearly.

Head components



Insect anatomy



Insect head parts. Legend: **a**, antennae; **c**, compound eye; **lb**, labium; **lr**, labrum; **md**, mandibles; **mx**, maxillae.

A typical insect head possesses a pair of antennae; eyes; mandibles, labrum, maxillae and labium (the latter four forming the cluster of "mouth parts", no. 32. in the diagram). Lying above the oesophagus is the brain or supraesophageal ganglion, divided into three pairs of ganglia: the protocerebrum, deutocerebrum and tritocerebrum from front to back (collectively no. 5 in the diagram). Nerves from the protocerebrum lead to the large compound eyes; from the deuterocerebrum to the antennae; and from the tritocerebrum to the labrum and stomatogastric nervous system. Circum-oesophageal connectives lead from the tritocerebrum around the gut to connect the brain to the ventral ganglionated nerve cord: nerves from the first three pairs of ganglia lead to the mandibles, maxillae and labium, respectively. The position of the mouth and the circum-oesophageal connectives allows a distinction to be made between pre- and post-oral structures; although it should be borne in mind that because structures can move around during development, a pre-oral position of a structure in the adult does not necessarily prove that its developmental origin is from there. The myriapod head is very similar to that of the insects.

The crustacean head is broadly similar to that of the insects, but possesses, in addition, a second pair of antennae that are innervated from the tritocerebrum. In place of the labium, crustaceans possess a second pair of maxillae.

Chelicerate head structures differ considerably from those of mandibulates (i.e. insects, crustaceans and myriapods); they possess eyes and a single pair of grasping appendages innervated from the brain, plus a labrum-like structure. Behind the mouth lies another pair of mouthparts, the pedipalps, and behind them lie the series of walking limbs. In chelicerates, the leg-bearing segments are fused with the anterior segments to form a prosoma, so that in living arthropods a distinct head only exists in mandibulates.

The acron concept

The arthropod head problem has until recently been predicated on the Articulata theory, i.e. that the arthropods and annelids are close relatives. Although arthropods are essentially direct developers that do not possess a trochophore-like larva, the annelids do. During annelid metamorphosis, segments are added close to the posterior of the body, behind the mouth; whereas the brain is derived from the episphere or region in front of the mouth. Recognition of this led to the concept of a primary, non-segmental component of the body in annelids known as the acron being developed, from which the brain is ultimately derived. Because the arthropod and annelid heads, in the light of the Articulata theory, were assumed to be structurally homologous in some way, the arthropod head was also often considered to incorporate a non-segmental acronal component. Taking the homology between annelid and arthropod heads at face value, Swedish workers such as Hanström and Holmgren assumed that a large part of the arthropod head must correspond to the acron, a view followed later by several prominent American insect workers such as Butt and Snodgrass. They proposed that all pre-oral structures in insects were non-segmental, although such a view is at odds with the preoral position of apparently *bona fide* appendages such as the antennae. A less extreme set of theories propose that only the protocerebrum and associated structures should be considered to be acronal.

The view that the arthropod head must contain an acronal remnant has been shaken by the relatively recent revision of protostome phylogeny, which has dismantled the Articulata and placed the arthropods together with a group of unsegmented worms often referred to as the Cycloneuralia in the so-called Ecdysozoa. All members of the Ecdysozoa are direct developers without a trochophore, and the cycloneuralians have terminal mouths. As a result, the idea of the arthropods having inherited a preoral acron from their ancestors seems less likely.

Molecular development and the arthropod head problem

The study of how developmental genes are expressed during embryogenesis has become an important new tool in the last twenty years for understanding the structure and evolution of morphology. The arthropod head problem has been tackled in three main ways in this regard, first by using genetic segmental markers to probe the obscure region in front of the mouth, especially in insects; second by looking at Hox gene expression patterns to detect patterns of homology between different arthropods; and third, by studying gene expression in particular features (especially the labrum) to determine its appendiculate or other status. Because all arthropods have the same complement of nine Hox loci, the morphological diversification observed is caused by these genes being expressed at different areas and times.

Areas of agreement

It is widely agreed that the insect, myriapod and crustacean heads are very similar. The apparent lack of a second antenna in insects and myriapods is explained by the idea that this appendage has been lost, leaving an appendage-less segment known as the intercalary segment. Modern phylogenies do not in general support an insect-myriapod relationship, suggesting that the second antenna has been lost independently in each group, perhaps as a result of a convergent adaptation to life on land. Furthermore, there is general agreement that the mandibles, first maxillae and labium/second maxillae each represent a post-oral segment; and that the first antenna represents a preoral segment.

Areas of disagreement

Areas of disagreement can be grouped into three categories: the nature of the pre-antennal region in mandibulates; the nature of the labrum; and the relationship between the chelicerate and mandibulate anterior segments.

Nature of the preoral region

The degree to which the area in front of the mouth is segmented remains one of the major controversies in the arthropod head problem. As already mentioned, earlier workers often considered the entire pre-oral region to be "acronal" and thus nonsegmental. Modern workers universally accept that at least the deuterocerebrum is segmental. However, the nature of the region in front of this is much less certain. Some molecular development studies have given limited support to the idea of an "ocular" segment corresponding to the protocerebrum; but these data are not unequivocal. The idea of the protocerebrum actually comprising two segments has also received support from both molecular and embryological data.

The labrum

The labrum is a flap-like structure that lies immediately in front of the mouth in almost all extant euarthropods, the general exception being provided by the probable chelicerate-relatives the pycnogonids. It has proved to be by far the most controversial of all arthropod head structures. It is innervated in crustaceans and insects from the tritocerebrum, i.e. the back of the brain. However, in development it often appears at the anterior of the head, and migrates backwards towards its adult position. Furthermore, it often appears as a bilobed structure, with a set of muscles, nerves and gene expression in many ways similar to that of an appendage. This evidence has been used to suggest that the labrum is in fact a highly reduced appendage. Its innervation from the rear of the brain has suggested to some workers that, if an appendage, it is the appendage of the tritocerebral segment; a point disputed by others who argue that the presence of a well-developed appendage in at least crustaceans in this segment (i.e., the second antenna, corresponding to the intercalary segment of insects) rules this out. If the labrum is an appendage then, it seems possible that its origin is indicated by its developmentally anterior position, i.e., that it is the appendage of a segment anterior to the first antenna. The most obvious choice for this is the segment whose ganglion is the protocerebrum, which in extant euarthropods bears no appendage (apart from the eyes). If the labrum is really an anterior appendage that has migrated to the posterior, then it may be homologous to the "antennae" of onychophorans which, as discussed below, seem to be innervated from a very anterior part of the brain, i.e. in front of the eyes. It has even been suggested (e.g., by Roonwal) that the labrum belongs to an even more obscure segment that lies in front of the ocular one. Nevertheless, many workers continue to be highly skeptical about the appendiculate nature of the labrum, preferring to see it as it appears, i.e., as an outgrowth of the body wall just in front of the mouth.

Mandibulate/chelicerate head homologies

Given the disagreements about the structure of the insect head, on which most effort has been spent, it is no surprise that the potential homologies between it and other arthropods, notably the chelicerates, are also very controversial. From after the second world war to the 1980s a commonly accepted model of arthropod evolution was that the extant euarthropods were polyphyletic, i.e. the main lineages had evolved independently from soft-bodied, annelid-like ancestors, following the work of Tiegs and especially Sidnie Manton. In this view, most of the head structures would also be convergent, and thus there was no point looking for specific homologies between major groups. However, the monophyletic theory of arthropod origins has since decisively gained the upper hand, which raises the problem of head homology once more.

The classical view was that the chelicerae were homologous to the second antennae of crustaceans (i.e., they are innervated from the tritocerebrum), a view based partly on the fact that the chelicerae were innervated from the same ganglion that innervates the labrum, which is the tritocerebrum in crustaceans and insects. Given that there are apparently no appendages in front of the chelicerae, the implication was that the deuterocerebrum had been lost in chelicerates (the protocerebrum innervates the eyes in both groups in this view). Whilst this view still has its defenders (notably Collete and Jacques Bitsch), the alternative view that the chelicerae are innervated from the deuterocerebrum has gained ground, based on molecular development in mites and spiders, and neuroanatomy in *Limulus*. If this is the case, then chelicerates simply have no tritocerebrum, i.e. there is no third supraoesophageal ganglion of the brain; the segment corresponding to it would be the suboesophageal pedipalp one. Such a theory does not, however, immediately account for the same ganglionic innervation of the chelicerae and labrum, although one solution is simply to claim that the labrum itself is not homologous between mandibulates and chelicerates (the view, for example, of Dieter Waloszek and colleagues).

The heads of onychophorans and tardigrades

The brain of onychophorans has been recently re-investigated and shown to possess two unusual features. First, although the mouth is ventral, as is the case in euarthropods, it is innervated from three different places; the sides, the posterior, and by a nerve that originates dorsally, and passes anteriorly down to curve back to the front of the mouth. This set of innervation makes sense if the mouth of onychophorans was originally terminal and has been bent downwards. Second, the antennae of the onychophorans appear to be innervated from in front of the eyes; which in euarthropod terms implies a protocerebral (or potentially even more anterior) innervation. This is supported by gene expression data, which show that the jaws too are derived from a protocerebral or dueterocerebral segment. As all euarthropod antennae are deuterocerebral or tritocerebral, this implies that the onychophoran antennae are not homologous to any euarthropod ones.

The tritocerebrum in arthropods is homologous to the third head segment in onychophora, which bears the slime glands (a pair of highly-modified appendages).

Fossil evidence

The Cambrian fossil record, above all the various lagerstätten such as the Burgess Shale, Sirius Passet, Chengjiang and Orsten faunas, has yielded a very rich record of well-preserved arthropods, including the well-known trilobites.

Many Cambrian arthropods, including the trilobites themselves, possess a single pair of slender antennae, which have been equated with either the first or second antennae of the crustaceans; and either the chelicerae of the missing appendages of the supposedly reduced deuterocerebrum in chelicerates. However, another group of arthropods, the so-called "great appendage" arthropods, including *Yohoia*, *Leanchoilia* and *Alalcomenaeus*, do not possess simple antennae, but rather have a robust, branched structure, which was called the "great appendage" by Harry B. Whittington in his restudy of these taxa. Yet another group of arthropods may possess two differentiated head appendages, of which the most important and controversial is the Chengjiang form *Fuxianhuia*. *Fuxianhuia* was claimed to possess a pair of short antennae anterior, followed by a robust pair of "sub-chelate" appendages. However, this assessment has been both disputed by Waloszek and colleagues, who consider that the sub-chelate appendages are in fact gut diverticulae; and supported by Graham Budd. Thus, its nature remains controverted at present. Other taxa have also been claimed to have a somewhat similar anterior appendage arrangement (e.g. *Fortiforceps*) but, with the exception of the well-preserved *Branchiocaris* from the Burgess Shale, most of them are highly equivocal.

In almost all Cambrian arthropods, the post-oral limbs show very little differentiation compared to the trunk limbs; the heads posterior to the mouth shows a considerable degree of variability, however, in the number of segments incorporated into the head.

Trilobites, in particular, possess a ventral sclerotised plate in the head called the hypostome. Whether this is homologous to the labrum or not is debated; although Waloszek and others have argued that as the phosphatocopines (upper stem-group crustaceans) seem to possess both, it cannot be.

Theories of Cambrian arthropod head segmentation

There are at least four main theories to account for anterior head appendages in Cambrian arthropods:

Scholtz and Edgecombe

Gerhard Scholtz and Greg Edgecombe accept that the antennae of onychophorans are protocerebral, and call them "primary" antennae to distinguish them from the "secondary" antennae of groups such as the insects and crustaceans. They also accept that taxa such as *Fuxianhuia* possess both antennae and "great appendages". Because in *Fuxianhuia* the antennae lie anterior to the great appendages, they suggest that these antennae are the inherited primitive "primary" antennae; and that the great appendages are thus equivalent to the first antennae of crustaceans. Because the secondary antennae are not present in stem group arthropods such as *Fuxianhuia*, nor in the extant chelicerates, they propose that arthropods, such as the trilobites, that possess secondary antennae, belong in a monophyletic group that also includes the mandibulates, called the Antennata. The trilobites are thus, in their view, not stem-group chelicerates, a commonly-held view, but rather, stem-group mandibulates. The status of the labrum is not resolved by this theory, but they argue that it the evidence for it being appendiculate is not compelling; thus it does not have to correspond to a well-developed appendage of any Cambrian arthropod.

Budd

Graham Budd's theory agrees with that of Scholtz and Edgecombe in accepting the protocerebral nature of the onychophoran antennae, and the two preoral appendages of Fuxianhuia. However, he traces the origin of the "great appendages" in the differentiated frontal appendages of Cambrian lobopods such as *Aysheaia* and *Kerygmachela*, neither of which possess convincing antennae. Thus, in Budd's view, the order of the two anterior appendages of taxa such as *Fuxianhuia* are reversed: the antennae are the first antennae (deutocerebral) of the mandibulates; and the great appendages correspond to the primary antennae of the onychophorans and Cambrian lobopods. Following previous work by Dewel and colleagues, Budd accounts for their reversal by arguing that the mouth in basal lobopods was terminal, and that as it rotated backwards and downwards, it brought the anterior appendage backwards with it. Given this transformation, it is likely, under this theory, that the remnant of the great appendage/primary antenna is the labrum of extant arthropods. Because in this view *Fuxianhuia* possesses both a hypostome and a great appendage, the hypostome cannot be straightforwardly homologous with the labrum.

Pycnogonids and the great appendage theory

Maxmen and others recently published a morphologically-based paper that claimed the enigmatic chelifores of extant pycnogonids are innervated from the protocerebrum, and not from the trito- or deutocerebrum as previously claimed. This would suggest that pycnogonids had uniquely retained a "great appendage" homologue as an appendage, unlike all other euarthropods in which it had been transformed into the labrum (pycnogonids lack a labrum). However, expression data of Hox genes that were published shortly afterwards suggested that the chelifores were deuterocerebral and thus most likely to be homologous to the chelicerae. The pycnogonids are thus neutral with regard to the great appendage theory.

Waloszek

Dieter Waloszek and colleagues have offered a rather different account of Cambrian arthropod head structure. They do not necessarily accept the primary antenna theory of the onychophoran antennae; and they reject the idea that *Fuxianhuia* or any of its close relatives possessed a great appendage. Rather, they place the "great appendage" arthropods in the stem-group of the chelicerates, arguing that the great appendage is homologous to the chelicerae of chelicerates, and the first antennae of crustaceans.

Cotton and Braddy

Trevor Cotton and Simon Braddy, in a comprehensive cladistic analysis of Cambrian arthropods, also proposed that the great appendage arthropods were stem-group chelicerates;and, accepting that *Fuxianhuia* and relatives possessed two preoral appendages, defended the classical view that the great appendage and the chelicerae were tritocerebral in origin; i.e. that the antennae of *Fuxianhuia* were deuterocerebral.

Assessment

The number and nature of the post-oral segments in the insect head have rarely been questioned. A much more difficult area, however, has been the nature of the preoral region. The obvious contradiction between a theory that no-preoral structures are segmental, and evidence, such as for the first antennae of crustaceans, that some such structures clearly are, led workers as long ago as Lankester to posit that there has been forward migration of segments in front of the mouth. Indeed, such a process can be seen in ontogeny of the tritocerebrum, which can be seen to migrate forward as the brain develops; furthermore, although in most insects and crustaceans its ganglia are part of the brain, its commissures still loop behind it, suggesting derivation from a more posterior position.

Nevertheless, even allowing for this possibility, the complexity of the anterior part of the brain, which even if the acron concept is incorrect may still have been inherited from very basal animals; untangling the new characters evolved by the earliest arthropods from those inherited from their ancestors therefore still stands centrally in the arthropod head problem.

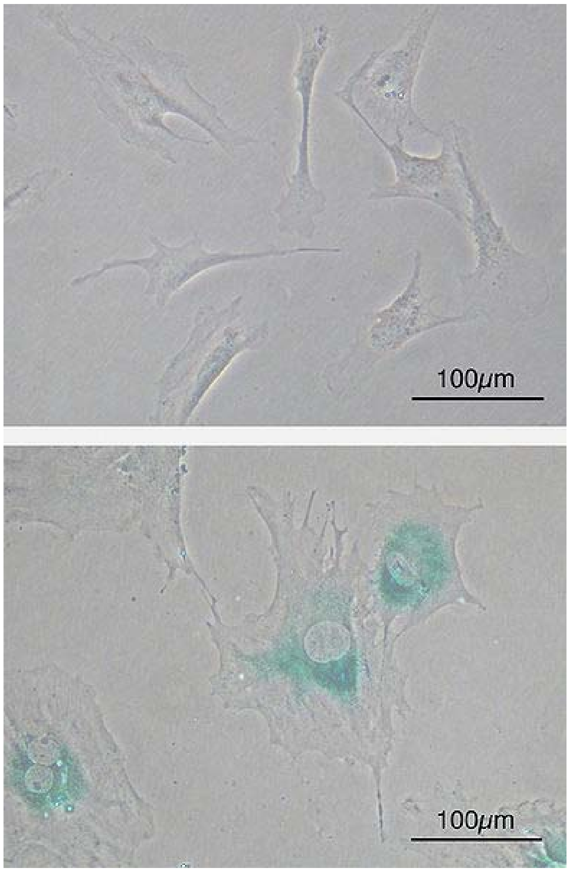
Chapter 2

Senescence

Senescence or **biological aging** is the change in the biology of an organism as it ages after its maturity. Such changes range from those affecting its cells and their function to that of the whole organism. There are a number of theories as to why senescence occurs, including ones that claim it is programmed by gene expression changes and that it is the accumulative damage of biological processes.

The word *senescence* is derived from the Latin word *senex*, meaning *old man*, *old age*, or *advanced in age*.

Cellular senescence



Cellular senescence
(upper) Primary mouse embryonic fibroblast cells (MEFs) before senescence. Spindle-shaped. (lower) MEFs became senescent after passages. Cells grow larger, flatten shape and expressed senescence-associated β-galactosidase (SABG, blue areas), a marker of cellular senescence.

Cellular senescence is the phenomenon by which normal diploid cells lose the ability to divide, normally after about 50 cell divisions in vitro. Some cells become senescent after fewer replications cycles as a result of DNA double strand breaks, toxins, etc. This phenomenon is also known as "replicative senescence", the "Hayflick phenomenon", or the Hayflick limit in honour of Dr. Leonard Hayflick who was the first to publish this information in 1965. In response to DNA damage (including shortened telomeres), cells either age or self-destruct (apoptosis, programmed cell death) if the damage cannot be easily repaired. In this 'cellular suicide', the death of one cell, or more, may benefit the organism as a whole. For example, in plants the death of the water-conducting xylem cells (tracheids and vessel elements) allows the cells to function more efficiently and so deliver water to the upper parts of a plant.

Aging of the whole organism

Organismal senescence is the aging of whole organisms. In general, aging is characterized by the declining ability to respond to stress, increased homeostatic imbalance, and increased risk of aging-associated diseases. Death is the ultimate consequence of aging, though "old age" is not a scientifically recognized cause of death because there is always a specific proximal cause, such as cancer, heart disease, or liver failure. Aging of whole organisms is therefore a complex process that can be defined as "a progressive deterioration of physiological function, an intrinsic age-related process of loss of viability and increase in vulnerability".

Differences in maximum life span among species correspond to different "rates of aging". For example, inherited differences in the rate of aging make a mouse elderly at 3 years and a human elderly at 80 years. These genetic differences affect a variety of physiological processes, including the efficiency of DNA repair, antioxidant enzymes, and rates of free radical production.

Supercentenarian Ann Pouder (8 April 1807 – 10 July 1917) photographed on her 110th birthday. A heavily lined face is common in human senescence.

Senescence of the organism gives rise to the Gompertz–Makeham law of mortality, which says that mortality rate rises rapidly with age.

Some animals, such as some reptiles and fish, age slowly (negligible senescence) and exhibit very long lifespans. Some even exhibit "negative senescence", in which mortality falls with age, in disagreement with the Gompertz–Makeham "law".

Whether replicative senescence (Hayflick limit) plays a causative role in organismal aging is at present an active area of investigation.

Theories of aging

The process of senescence is complex, and may derive from a variety of different mechanisms and exist for a variety of different reasons. However, senescence is not universal, and scientific evidence suggests that cellular senescence evolved in certain species because it prevents the onset of cancer. In a few simple species, such as Hydra, senescence is negligible and cannot be detected. All such species have no "post-mitotic" cells; they reduce the effect of damaging free radicals by cell division and dilution. Such species are not immortal, however, as they will eventually fall prey to trauma or disease. Moreover, average lifespans can vary greatly within and between species. This suggests that both genetic and environmental factors contribute to aging.

In general, theories that explain senescence have been divided between the programmed and stochastic theories of aging. Programmed theories imply that aging is regulated by biological clocks operating throughout the lifespan. This regulation would depend on changes in gene expression that affect the systems responsible for maintenance, repair, and defense responses. Stochastic theories blame environmental impacts on living organisms that induce cumulative damage at various levels as the cause of aging, examples of which ranging from damage to DNA, damage to tissues and cells by oxygen radicals (widely known as free radicals countered by the even more well-known antioxidants), and cross-linking.

However, aging is seen as a progressive failure of homeodynamics (homeostasis) involving genes for the maintenance and repair, stochastic events leading to molecular damage and molecular heterogeneity, and chance events determining the probability of death. Since complex and interacting systems of maintenance and repair comprise the homeodynamic (old term: homeostasis) space of a biological system, aging is considered to be a progressive shrinkage of homeodynamic space mainly due to increased molecular heterogeneity.

Evolutionary theories

A gene can be expressed at various life-stages. Therefore, natural selection can support lethal and harmful alleles, if their expression occurs after reproduction. Senescence may be the product of such selection. In addition, aging is believed to have evolved because of the increasingly smaller probability of an organism still being alive at older age, due to predation and accidents, both of which may be random and age-invariant. It is thought that strategies that result in a higher reproductive rate at a young age, but shorter overall lifespan, result in a higher lifetime reproductive success and are therefore favoured by natural selection. In essence, aging is, therefore, the result of investing resources in reproduction, rather than maintenance of the body (the "Disposable Soma" theory), in light of the fact that accidents, predation, and disease will eventually kill the organism no matter how much energy is devoted to repair of the body. Various other theories of aging exist, and are not necessarily mutually exclusive.

The geneticist J. B. S. Haldane wondered why the dominant mutation that causes Huntington's disease remained in the population, and why natural selection had not eliminated it. The onset of this neurological disease is (on average) at age 45 and is invariably fatal within 10–20 years. Haldane assumed that, in human prehistory, few survived until age 45. Since few were alive at older ages and their contribution to the next generation was therefore small relative to the large cohorts of younger age groups, the force of selection against such late-acting deleterious mutations was correspondingly small. However, if a mutation affected younger individuals, selection against it would be strong. Therefore, late-acting deleterious mutations could accumulate in populations over evolutionary time through genetic drift. This principle has been demonstrated experimentally. And it is these later-acting deleterious mutations that are believed to cause—even allow—age-related mortality.

Peter Medawar formalised this observation in his mutation accumulation theory of aging. "The force of natural selection weakens with increasing age—even in a theoretically immortal population, provided only that it is exposed to real hazards of mortality. If a genetic disaster… happens late enough in individual life, its consequences may be completely unimportant". The 'real hazards of mortality' are, in typical circumstances, predation, disease, and accidents. So, even an immortal population, whose fertility does not decline with time, will have fewer individuals alive in older age groups. This is called 'extrinsic mortality'. Young cohorts, not depleted in numbers yet by extrinsic mortality, contribute far more to the next generation than the few remaining older cohorts, so the force of selection against late-acting deleterious mutations, which affect only these few older individuals, is very weak. The mutations may not be selected against, therefore, and may spread over evolutionary time into the population.

The major testable prediction made by this model is that species that have high extrinsic mortality in nature will age more quickly and have shorter intrinsic lifespans. This is borne out among mammals, the best-studied in terms of life history. There is a correlation among mammals between body size and lifespan, such that larger species live longer than smaller species in controlled/optimum conditions, but there are notable exceptions. For instance, many bats and rodents are of similar size, yet bats live much longer. For instance, the little brown bat, half the size of a mouse, can live 30 years in the wild. A mouse will live 2–3 years even with optimum conditions. The explanation is that bats have fewer predators, so therefore low extrinsic mortality. Thus, more individuals survive to later ages, so the force of selection against late-acting deleterious mutations is stronger. Fewer late-acting deleterious mutations = slower aging = longer lifespan. Birds are also warm-blooded and are of similar size to many small mammals, yet live often 5–10 times as long. They have fewer predation pressures compared with ground-dwelling mammals. And seabirds, which, in general, have the fewest predators of all birds, live longest.

Also, when examining the body-size vs. lifespan relationship, one will observe that predator mammals tend to have longer lifespans than prey animals in a controlled environment such as a zoo or nature reserve. The explanation for the long lifespans of primates (such as humans, monkeys, and apes) relative to body size is that their intelligence and often sociality helps them avoid becoming prey. Being a predator, being smart, and working together all reduce extrinsic mortality.

Another evolutionary theory of aging was proposed by George C. Williams and involves antagonistic pleiotropy. A single gene may affect multiple traits. Some traits that increase fitness early in life may also have negative effects later in life. But, because many more individuals are alive at young ages than at old ages, even small positive effects early can be strongly selected for, and large negative effects later may be very weakly selected against. Williams suggested the following example: Perhaps a gene codes for calcium deposition in bones, which promotes juvenile survival and will therefore be favored by natural selection; however, this same gene promotes calcium deposition in the arteries, causing negative effects in old age. Therefore, negative effects in old age may reflect the result of natural selection for pleiotropic genes that are beneficial early in life. In this case, fitness is relatively high when Fisher's reproductive value is high and relatively low when Fisher's reproductive value is low.

Gene regulation

A number of genetic components of aging have been identified using model organisms, ranging from the simple budding yeast *Saccharomyces cerevisiae* to worms such as *Caenorhabditis elegans* and fruit flies (*Drosophila melanogaster*). Study of these organisms has revealed the presence of at least two conserved aging pathways.

One of these pathways involves the gene *Sir2*, a NAD+-dependent histone deacetylase. In yeast, Sir2 is required for genomic silencing at three loci: The yeast mating loci, the telomeres and the ribosomal DNA (rDNA). In some species of yeast, replicative aging may be partially caused by homologous recombination between rDNA repeats; excision of rDNA repeats results in the formation of extrachromosomal rDNA circles (ERCs). These ERCs replicate and preferentially segregate to the mother cell during cell division, and are believed to result in cellular senescence by titrating away (competing for) essential nuclear factors. ERCs have not been observed in other species (nor even all strains of the same yeast species) of yeast (which also display replicative senescence), and ERCs are not believed to contribute to aging in higher organisms such as humans (they have not been shown to accumulate in mammals in a similar manner to yeast). Extrachromosomal circular DNA (eccDNA) has been found in worms, flies, and humans. The origin and role of eccDNA in aging, if any, is unknown.

Despite the lack of a connection between circular DNA and aging in higher organisms, extra copies of Sir2 are capable of extending the lifespan of both worms and flies (though, in flies, this finding has not been replicated by other investigators, and the activator of Sir2 resveratrol does not reproducibly increase lifespan in either species). Whether the Sir2 homologues in higher organisms have any role in lifespan is unclear, but the human SIRT1 protein has been demonstrated to deacetylate p53, Ku70, and the forkhead family of transcription factors. SIRT1 can also regulate acetylates such as CBP/p300, and has been shown to deacetylate specific histone residues.

RAS1 and RAS2 also affect aging in yeast and have a human homologue. RAS2 overexpression has been shown to extend lifespan in yeast.

Other genes regulate aging in yeast by increasing the resistance to oxidative stress. Superoxide dismutase, a protein that protects against the effects of mitochondrial free radicals, can extend yeast lifespan in stationary phase when overexpressed.

In higher organisms, aging is likely to be regulated in part through the insulin/IGF-1 pathway. Mutations that affect insulin-like signaling in worms, flies, and the growth hormone/IGF1 axis in mice are associated with extended lifespan. In yeast, Sir2 activity is regulated by the nicotinamidase PNC1. PNC1 is transcriptionally upregulated under stressful conditions such as caloric restriction, heat shock, and osmotic shock. By converting nicotinamide to niacin, nicotinamide is removed, inhibiting the activity of Sir2. A nicotinamidase found in humans, known as PBEF, may serve a similar function, and a secreted form of PBEF known as visfatin may help to regulate serum insulin levels. It is not known, however, whether these mechanisms also exist in humans, since there are obvious differences in biology between humans and model organisms.

Sir2 activity has been shown to increase under calorie restriction. Due to the lack of available glucose in the cells, more NAD+ is available and can activate Sir2. Resveratrol, a stilbenoid found in the skin of red grapes, was reported to extend the lifespan of yeast, worms, and flies (the lifespan extension in flies and worms have proved irreproducible by independent investigators). It has been shown to activate Sir2 and therefore mimics the effects of calorie restriction, if one accepts that caloric restriction is indeed dependent on Sir2.

Gene expression is imperfectly controlled, and it is possible that random fluctuations in the expression levels of many genes contribute to the aging process as suggested by a study of such genes in yeast. Individual cells, which are genetically identical, none-the-less can have substantially different responses to outside stimuli, and markedly different lifespans, indicating the epigenetic factors play an important role in gene expression and aging as well as genetic factors.

According to the GenAge database of aging-related genes there are over 700 genes associated with aging in model organisms: 555 in the soil roundworm (*Caenorhabditis elegans*), 87 in the bakers' yeast (*Saccharomyces cerevisiae*), 75 in the fruit fly (*Drosophila melanogaster*) and 68 in the mouse (*Mus musculus*). The following is a list of genes connected to longevity through research on model organisms:

Podospora	**Saccharomyces**	**Caenorhabditis**	**Drosophila**	**Mus**
grisea	LAG1	daf-2	sod1	Prop-1
	LAC1	age-1/daf-23	cat1	p66shc (Not independently verified)
	pit-1	Ghr		
	RAS1	daf-18	mth	mclk1
	RAS2	akt-1/akt-2		
	PHB1	daf-16		
	PHB2	daf-12		
	CDC7	ctl-1		
	BUD1	old-1		

RTG2	spe-26
RPD3	clk-1
HDA1	mev-1
SIR2	
	aak-2
SIR4-42	
UTH4	
YGL023	
SGS1	
RAD52	
FOB1	

Cellular senescence

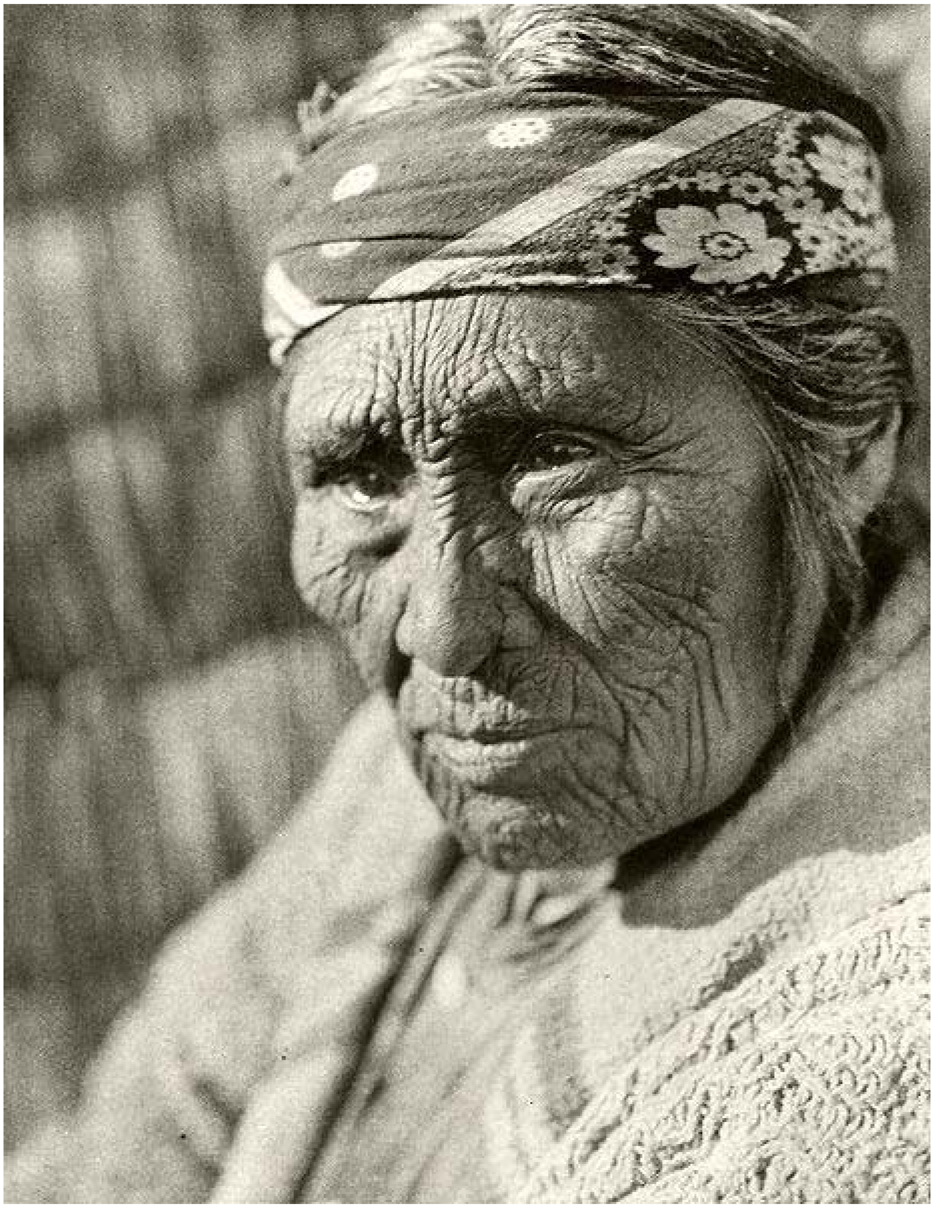
As noted above, senescence is not universal, and senescence is not observed in single-celled organisms that reproduce through the process of cellular mitosis. Moreover, cellular senescence is not observed in several organisms, including perennial plants, sponges, corals, and lobsters. In those species where cellular senescence is observed, cells eventually become post-mitotic when they can no longer replicate themselves through the process of cellular mitosis; i.e., cells experience *replicative senescence*. How and why some cells become post-mitotic in some species has been the subject of much research and speculation, but (as noted above) it is widely believed that cellular senescence evolved as a way to prevent the onset and spread of cancer. Somatic cells that have divided many times will have accumulated DNA mutations and would therefore be in danger of becoming cancerous if cell division continued.

Lately, the role of telomeres in cellular senescence has aroused general interest, especially with a view to the possible genetically adverse effects of cloning. The successive shortening of the chromosomal telomeres with each cell cycle is also believed to limit the number of divisions of the cell, thus contributing to aging. There have, on the other hand, also been reports that cloning could alter the shortening of telomeres. Some cells do not age and are, therefore, described as being "biologically immortal". It is theorized by some that when it is discovered exactly what allows these cells, whether it be the result of telomere lengthening or not, to divide without limit that it will be possible to genetically alter other cells to have the same capability. It is further theorized that it will eventually be possible to genetically engineer all cells in the human body to have this capability by employing gene therapy and, therefore, stop or reverse aging, effectively making the entire organism potentially immortal.

Cancer cells are usually immortal. In about 85% of tumors, this evasion of cellular senescence is the result of up-activation of their telomerase genes. This simple observation suggests that reactivation of telomerase in healthy individuals could greatly increase their cancer risk.

Whether cell senescence plays any role in organismal aging is at present unknown, and is an active area of investigation. Mouse mutants lacking telomerase do not immediately show accelerated aging.

Chemical damage



Elderly Klamath woman photographed by Edward S. Curtis in 1924

One of the earliest aging theories was the *Rate of Living Hypothesis* described by Raymond Pearl in 1928(based on earlier work by Max Rubner), which states that fast basal metabolic rate corresponds to short maximum life span.

While there may be some validity to the idea that for various types of specific damage detailed below that are by-products of metabolism, all other things being equal, a fast metabolism may reduce lifespan, in general this theory does not adequately explain the differences in lifespan either within, or between, species. Calorically-restricted animals process as much, or more, calories per gram of body mass, as their *ad libitum* fed counterparts, yet exhibit substantially longer lifespans. . Similarly, metabolic rate is a poor predictor of lifespan for birds, bats and other species that, it is presumed, have reduced mortality from predation, and therefore have evolved long lifespans even in the presence of very high metabolic rates. More recently, it was shown that, when modern statistical methods for correcting for the effects of body size and phylogeny are employed, metabolic rate does not correlate with longevity in mammals or birds.

With respect to specific types of chemical damage caused by metabolism, it is suggested that damage to long-lived biopolymers, such as structural proteins or DNA, caused by ubiquitous chemical agents in the body such as oxygen and sugars, are in part responsible for aging. The damage can include breakage of biopolymer chains, cross-linking of biopolymers, or chemical attachment of unnatural substituents (haptens) to biopolymers.

Under normal aerobic conditions, approximately 4% of the oxygen metabolized by mitochondria is converted to superoxide ion, which can subsequently be converted to hydrogen peroxide, hydroxyl radical and eventually other reactive species including other peroxides and singlet oxygen, which can, in turn, generate free radicals capable of damaging structural proteins and DNA. Certain metal ions found in the body, such as copper and iron, may participate in the process. (In Wilson's disease, a hereditary defect that causes the body to retain copper, some of the symptoms resemble accelerated senescence.) These processes are termed *oxidative damage* and are linked to the benefits of nutritionally derived polyphenol antioxidants.

Sugars such as glucose and fructose can react with certain amino acids such as lysine and arginine and certain DNA bases such as guanine to produce sugar adducts, in a process called *glycation*. These adducts can further rearrange to form reactive species, which can then cross-link the structural proteins or DNA to similar biopolymers or other biomolecules such as non-structural proteins. People with diabetes, who have elevated blood sugar, develop senescence-associated disorders much earlier than the general population, but can delay such disorders by rigorous control of their blood sugar levels. There is evidence that sugar damage is linked to oxidant damage in a process termed *glycoxidation*.

Free radicals can damage proteins, lipids or DNA. Glycation mainly damages proteins. Damaged proteins and lipids accumulate in lysosomes as lipofuscin. Chemical damage to structural proteins can lead to loss of function; for example, damage to collagen of blood vessel walls can lead to vessel-wall stiffness and, thus, hypertension, and vessel wall thickening and reactive tissue formation (atherosclerosis); similar processes in the kidney can lead to renal failure. Damage to enzymes reduces cellular functionality. Lipid peroxidation of the inner mitochondrial membrane reduces the electric potential and the ability to generate energy. It is probably no accident that nearly all of the so-called "accelerated aging diseases" are due to defective DNA repair enzymes.

It is believed that the impact of alcohol on aging can be partly explained by alcohol's activation of the HPA axis, which stimulates glucocorticoid secretion, long-term exposure to which produces symptoms of aging.

Reliability theory

Reliability theory suggests that biological systems start their adult life with a high load of initial damage. Reliability theory is a general theory about systems failure. It allows researchers to predict the age-related failure kinetics for a system of given architecture (reliability structure) and given reliability of its components. Reliability theory predicts that even those systems that composed entirely of non-aging elements (with a constant failure rate) will nevertheless deteriorate (fail more often) with age, if these systems are redundant in irreplaceable elements. Aging, therefore, is a direct consequence of systems.

Reliability theory also predicts the late-life mortality deceleration with subsequent leveling-off, as well as the late-life mortality plateaus, as an inevitable consequence of redundancy exhaustion at extreme old ages. The theory explains why mortality rates increase exponentially with age (the Gompertz law) in many species, by taking into account the initial flaws (defects) in newly formed systems. It also explains why organisms "prefer" to die according to the Gompertz law, while technical devices usually fail according to the Weibull (power) law. Reliability theory allows to specify conditions when organisms die according to the Weibull distribution: Organisms should be relatively free of initial flaws and defects. The theory makes it possible to find a general failure law applicable to all adult and extreme old ages, where the Gompertz and the Weibull laws are just special cases of this more general failure law. The theory explains why relative differences in mortality rates of compared populations (within a given species) vanish with age (compensation law of mortality), and mortality convergence is observed due to the exhaustion of initial differences in redundancy levels.

Miscellaneous

Recently, a kind of early senescence has been alleged to be a possible unintended outcome of early cloning experiments. The issue was raised in the case of Dolly the sheep, following her death from a contagious lung disease. The claim that Dolly's early death involved premature senescence has been vigorously contested, and Dolly's creator, Dr. Ian Wilmut has expressed the view that her illness and death were probably unrelated to the fact that she was a clone.

A set of rare hereditary (genetic) disorders, each called progeria, has been known for some time. Sufferers exhibit symptoms resembling accelerated aging, including wrinkled skin. The cause of Hutchinson–Gilford progeria syndrome was reported in the journal *Nature* in May 2003. This report suggests that DNA damage, not oxidative stress, is the cause of this form of accelerated aging.

Chapter 3

Extraterrestrial Life

Extraterrestrial life (from the Latin words: *extra* ("beyond", or "not of") and *terrestris* ("of or belonging to Earth")) is defined as life that does not originate from Earth. Hypothetical forms of extraterrestrial life range from simple bacteria-like organisms to sapient beings far more advanced than humans. It is currently unknown whether any such forms of life exist or ever existed.

The development and testing of theories about extraterrestrial life is known as exobiology or astrobiology; the term astrobiology however also covers the study of life on Earth, viewed in its astronomical context.

Background

Various controversial claims have been made for evidence of extraterrestrial life. A less direct argument for the existence of extraterrestrial life relies on the vast size of the observable Universe. According to this argument, supported by scientists such as Carl Sagan and Stephen Hawking, it would be improbable for life *not* to exist somewhere other than Earth.

One possibility is that life has emerged independently at many places throughout the Universe. Another possibility is panspermia or exogenesis, in which life would have spread between habitable planets. These two hypotheses are not necessarily mutually exclusive.

Suggested locations at which life might have developed, or which might continue to host life today, include the planets Venus and Mars; moons of Jupiter, such as Europa; moons of Saturn, such as Titan and Enceladus; and extrasolar planets, such as Gliese 581 c, g and d, recently discovered to be near Earth mass and apparently located in their star's habitable zone, with the potential to have liquid water.

Beliefs that some unidentified flying objects are of extraterrestrial origin, along with claims of alien abduction, are considered bogus by most scientists. Most UFO sightings are explained either as sightings of Earth-based aircraft or known astronomical objects, or as hoaxes.

Possible basis of extraterrestrial life

Several theories have been proposed about the possible basis of alien life from a biochemical, evolutionary or morphological viewpoint.

Alien life, such as bacteria, has been theorized by scientists such as Carl Sagan to exist in the Solar System and quite possibly throughout the Universe. No samples have been found.

Biochemistry

All life on Earth requires carbon, hydrogen, nitrogen, oxygen, phosphorus and sulfur (CHNOPS) as well as numerous other elements in smaller amounts. Life also requires water as the solvent in which biochemical reactions take place. Sufficient quantities of carbon and the other major life-forming elements, along with water, may enable the formation of living organisms on other planets with a chemical make-up and average temperature similar to that of Earth. Because Earth and other planets are made up of "stardust", i.e. relatively abundant chemical elements formed from stars which have ended their lives as supernovae, it is very probable that other planets may have been formed by elements of a similar composition to the Earth's. The combination of carbon, hydrogen and oxygen in the chemical form of carbohydrates (e.g. sugar) can be a source of chemical energy on which life depends, and can also provide structural elements for life (such as ribose, in the molecules DNA and RNA, and cellulose in plants). Plants derive energy through the conversion of light energy into chemical energy via photosynthesis. Life, as currently recognized, requires carbon in both reduced (methane derivatives) and partially-oxidized (carbon oxides) states. It also appears to require nitrogen as a reduced ammonia derivative in all proteins, sulfur as a derivative of hydrogen sulfide in some necessary proteins, and phosphorus oxidized to phosphates in genetic material and in energy transfer. Adequate water as a solvent supplies adequate oxygen as constituents of biochemical substances.

Pure water is useful because it has a neutral pH due to its continued dissociation between hydroxide and hydronium ions. As a result, it can dissolve both positive metallic ions and negative non-metallic ions with equal ability. Furthermore, the fact that organic molecules can be either hydrophobic (repelled by water) or hydrophilic (soluble in water) creates the ability of organic compounds to orient themselves to form water-enclosing membranes. The fact that solid water (ice) is less dense than liquid water (within specific temperature ranges) also means that ice floats, thereby preventing Earth's oceans from slowly freezing. Without this quality, the oceans could have frozen solid during the Snowball Earth episodes. Additionally, the hydrogen bonds between water molecules give it an ability to store energy with evaporation, which upon condensation is released. This helps to moderate the climate, cooling the tropics and warming the poles, helping to maintain the thermodynamic stability needed for life.

Carbon is fundamental to terrestrial life for its immense flexibility in creating covalent chemical bonds with a variety of non-metallic elements, principally nitrogen, oxygen and hydrogen. Carbon dioxide and water together enable the storage of solar energy in sugars, such as glucose. The oxidation of glucose releases biochemical energy needed to fuel all other biochemical reactions.

The ability to form organic acids (–COOH) and amine bases (–NH_2) gives rise to the possibility of neutralization dehydrating reactions to build long polymer peptides and catalytic proteins from monomer amino acids, and with phosphates to build not only DNA (the information-storing molecule of inheritance), but also ATP (the principal energy "currency" of cellular life).

Due to their relative abundance and usefulness in sustaining life, many have hypothesized that life forms elsewhere in the universe would also utilize these basic materials. However, other elements and solvents could also provide a basis for life. Silicon is most often deemed to be the probable alternative to carbon. Silicon life forms are proposed to have a crystalline morphology, and are theorized to be able to exist in high temperatures, such as on planets which are very close to their star. Life forms based in ammonia (rather than water) have also been suggested, though this solution appears less optimal than water.

From a chemical perspective, life is fundamentally a self-replicating reaction, but one which could arise under a great many conditions and with various possible ingredients, though carbon-oxygen within the liquid temperature range of water seems most conducive. Suggestions have even been made that self-replicating reactions of some sort could occur within the plasma of a star, though it would be highly unconventional.

Several pre-conceived ideas about the characteristics of life outside of Earth have been questioned. For example, NASA scientists believe that the color of photosynthesizing pigments on extrasolar planets might not be green.

Evolution and morphology

In addition to the biochemical basis of extraterrestrial life, many have also considered evolution and morphology. Science fiction has often depicted extraterrestrial life with humanoid and/or reptilian forms. Aliens have often been depicted as having light green or grey skin, with a large head, as well as four limbs—i.e. fundamentally humanoid. Other subjects, such as felines and insects, etc., have also occurred in fictional representations of aliens.

A division has been suggested between universal and parochial (narrowly restricted) characteristics. Universals are features which are thought to have evolved independently more than once on Earth (and thus, presumably, are not too difficult to develop) and are so intrinsically useful that species will inevitably tend towards them. The most fundamental of these is probably bilateral symmetry, but more complex (though still basic) characteristics include flight, sight, photosynthesis and limbs, all of which are thought to have evolved several times here on Earth. There is a huge variety of eyes, for example, and many of these have radically different working schematics and different visual foci: the visual spectrum, infrared, polarity and echolocation. Parochials, however, are essentially arbitrary evolutionary forms. These often have little inherent utility (or at least have a function which can be equally served by dissimilar morphology) and probably will not be replicated. Intelligent aliens could communicate through gestures, as deaf humans do, or by sounds created from structures unrelated to breathing, which happens on Earth when, for instance, cicadas vibrate their wings, or crickets rub their legs.

Attempting to define parochial features challenges many taken-for-granted notions about morphological necessity. Skeletons, which are essential to large terrestrial organisms according to the experts of the field of gravitational biology, are almost assured to be replicated elsewhere in one form or another. The assumption of radical diversity amongst putative extraterrestrials is by no means settled. While many exobiologists do stress that the enormously heterogeneous nature of life on Earth foregrounds an even greater variety in outer space, others point out that convergent evolution may dictate substantial similarities between Earth and extraterrestrial life. These two schools of thought are called "divergionism" and "convergionism" respectively.

Beliefs in extraterrestrial life

Ancient and medieval ideas

In antiquity, it was common to assume a cosmos consisting of "many worlds" inhabited by intelligent, non-human life-forms, but these "worlds" were mythological and not informed by the heliocentric understanding of the solar system, or the understanding of the Sun as one among countless stars. An example would be the fourteen loka of Hindu cosmology, or the Nine Worlds of Old Norse mythology, etc. The Sun and the Moon often appear as inhabited worlds in such contexts, or as vehicles (chariots or boats, etc.) driven by gods. The Japanese folk tale of *The Tale of the Bamboo Cutter* (10th century) is an example of a princess of the Moon people visiting Earth.

Such conceptions of a universe consisting of "many worlds" are also found in classical Greek philosophy, and later in Christian and Jewish theology. The first important thinkers to argue systematically for a Universe full of other planets and, therefore, possible extraterrestrial life were the ancient Greek writer Thales and his student Anaximander in the 7th and 6th centuries B.C. The atomists of Greece like Epicurus took up the idea, arguing that an infinite universe ought to have an infinity of populated worlds. Ancient Greek cosmology worked against the idea of extraterrestrial life in one critical respect, however: the geocentric Universe. Championed by Aristotle and codified by Ptolemy, it favored the Earth and Earth-life (Aristotle denied that there could be a plurality of worlds) and seemingly rendered extraterrestrial life philosophically untenable. Lucian of Samosata, in his novels, described inhabitants of the Moon and other celestial bodies as humanoids, but significantly different from humans.

The Jewish Talmud states that there are at least 18,000 other worlds, but provides little elaboration on the nature of those worlds, or on whether they are physical or spiritual.

Based on this, however, the 18th century exposition "Sefer HaB'rit" posits that extraterrestrial creatures exist, and that some may well possess intelligence. It adds that humans should not expect creatures from another world to resemble earthly life any more than sea creatures resemble land animals.

Hindu beliefs of endlessly repeated cycles of life have led to descriptions of multiple worlds in existence . According to Hindu scriptures, there are innumerable universes to facilitate the fulfillment of the separated desires of innumerable living entities. However, the purpose of such creations is to bring back the deluded souls to correct understanding about the purpose of life. Aside from the innumerable universes which are material, there is also the unlimited spiritual world, where the purified living entities live with perfect conception about life and ultimate reality. The spiritually aspiring saints and devotees, as well as thoughtful men of the material world, have been getting guidance and help from these purified living entities of the spiritual world from time immemorial. However, the relevance of such descriptions has to be evaluated in the context of a correct understanding of geography and science at those times.

Within Islam, the statement of the Qur'an "All praise belongs to God, Lord of all the worlds" indicates multiple universal bodies, and maybe even multiple universes, which may indicate extraterrestrial and even extradimensional life. Surat Al-Jinn also mentions a statement from a Jinn regarding the current status and ability of his group in the heavens.

According to Ahmadiyya Islam a more direct reference from the Quran is presented by Mirza Tahir Ahmad as a proof that life on other planets may exist according to the Quran. In his book, *Revelation, Rationality, Knowledge & Truth*, he quotes verse 42:29 "And among His Signs is the creation of the heavens and the earth, and of whatever living creatures (*da'bbah*) He has spread forth in both..."; according to this verse there is life in heavens. According to the same verse "And He has the power to gather them together (*jam-'i-him*) when He will so please"; indicates the bringing together the life on Earth and the life elsewhere in the Universe. The verse does not specify the time or the place of this meeting but rather states that this event will most certainly come to pass whenever God so desires. It should be pointed out that the Arabic term Jam-i-him used to express the gathering event can imply either a physical encounter or a contact through communication.

In Shia Islam the 6th Imam Ja'far al-Sadiq has been quoted as saying that there are living beings on other planets. He has also stated that they may be more intelligent or advanced than humans.

When Christianity spread throughout the West, the Ptolemaic system became very widely accepted, and although the Church never issued any formal pronouncement on the question of alien life, at least tacitly, the idea was aberrant. In 1277, the Bishop of Paris, Étienne Tempier, did overturn Aristotle on one point: God *could* have created more than one world (given His omnipotence). Taking a further step, and arguing that aliens actually existed, remained rare. Notably, Cardinal Nicholas of Kues speculated about aliens on the Moon and Sun.

Early modern period

Giordano Bruno, *De l'Infinito, Universo e Mondi, 1584*

There was a dramatic shift in thinking initiated by the invention of the telescope and the Copernican assault on geocentric cosmology. Once it became clear that the Earth was merely one planet amongst countless bodies in the universe, the extraterrestrial idea moved towards the scientific mainstream. The best known early-modern proponent of such ideas was the Italian philosopher Giordano Bruno, who argued in the 16th century for an infinite Universe in which every star is surrounded by its own planetary system. Bruno's thoughts about God and the Universe, of which many contradicted essential dogmas of the Catholic Faith, led to his eventual condemnation as a heretic by a tribunal of the Roman Inquisition of the Roman Catholic Church. Contemporary civil authorities enforced the penal statute of Emperor Frederick II Inconsutilem Tunicam of 12 February 1220 mandating the burning of heretics, resulting in his being among the last of heretics to be burned alive at the stake in Rome in the year 1600.

In the early 17th century, the Czech astronomer Anton Maria Schyrleus of Rheita mused that "if Jupiter has (...) inhabitants (...) they must be larger and more beautiful than the inhabitants of the Earth, in proportion to the [characteristics] of the two spheres". The Catholic Church has not made a formal ruling on the existence of extraterrestrials. However, writing in the Vatican newspaper, the astronomer, Father José Gabriel Funes, director of the Vatican Observatory near Rome, said in 2008 that intelligent beings created by God could exist in outer space.

In Baroque literature such as *The Other World: The Societies and Governments of the Moon* by Cyrano de Bergerac, extraterrestrial societies are presented as humoristic or ironic parodies of earthly society. The didactic poet Henry More took up the classical theme of the Greek Democritus in "Democritus Platonissans, or an Essay Upon the Infinity of Worlds" (1647). In "The Creation: a Philosophical Poem in Seven Books" (1712), Sir Richard Blackmore observed: "We may pronounce each orb sustains a race / Of living things adapted to the place". With the new relative viewpoint that the Copernican revolution had wrought, he suggested "our world's sunne / Becomes a starre elsewhere". Fontanelle's "Conversations on the Plurality of Worlds" (translated into English in 1686) offered similar excursions on the possibility of extraterrestrial life, expanding, rather than denying, the creative sphere of a Maker.

The possibility of extraterrestrials remained a widespread speculation as scientific discovery accelerated. William Herschel, the discoverer of Uranus, was one of many 18th–19th century astronomers convinced that the Solar System, and perhaps others, would be well-populated by alien life. Other luminaries of the period who championed "cosmic pluralism" included Immanuel Kant and Benjamin Franklin. At the height of the Enlightenment, even the Sun and Moon were considered candidates for extraterrestrial inhabitants.

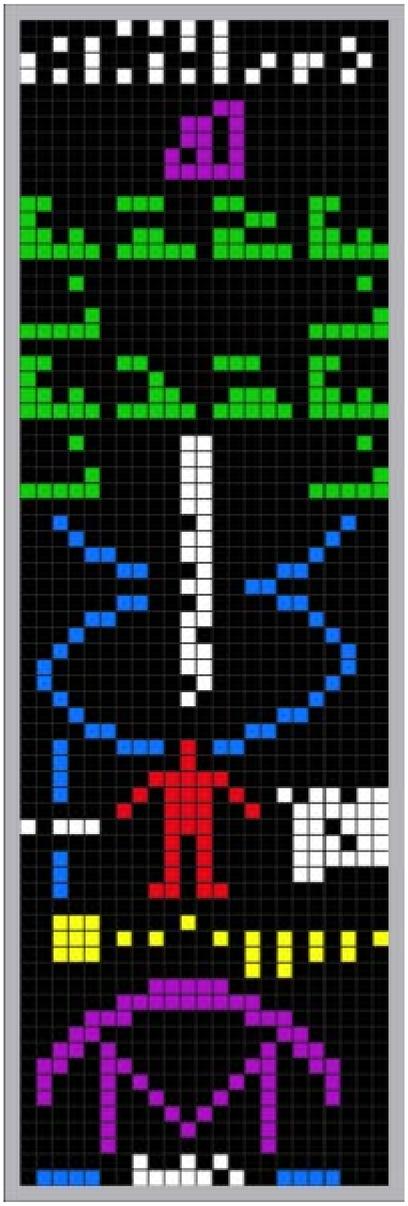
19th century

In 1854, William Whewell, a fellow of Trinity College, Cambridge, who popularized the word *scientist,* theorized that Mars had seas, land and possibly life forms. Speculation about life on Mars exploded in the late 19th century, following telescopic observation by some observers of apparent Martian canals — which were however soon found to be optical illusions. Despite this, in 1895, American astronomer Percival Lowell published his book *Mars,* followed by *Mars and its Canals* in 1906, proposing that the canals were the work of a long-gone civilization. This idea led British writer H. G. Wells to write *The War of the Worlds* in 1897, telling of an invasion by aliens from Mars who were fleeing the planet’s desiccation.

Spectroscopic analysis of Mars' atmosphere began in earnest in 1894, when U.S. astronomer William Wallace Campbell showed that neither water nor oxygen were present in the Martian atmosphere. By 1909 better telescopes and the best perihelic opposition of Mars since 1877 conclusively put an end to the canal theory.

The science fiction genre, although not so named during the time, develops during the late 19th century. Jules Verne's *Around the Moon* (1870) features a discussion of the possibility of life on the Moon, but with the conclusion that it is barren. Stories involving extraterrestrials are found in e.g. Garrett P. Serviss's *Edison's Conquest of Mars* (1897). *The War of the Worlds* by H. G. Wells was published in 1898 and stands at the beginning of the popular idea of the "Martian invasion" of Earth prominent in 20th century pop culture.

20th century



The Arecibo message is a digital message sent to globular star cluster M13, and is a well-known symbol of human attempts to contact extraterrestrials.

A radio drama version of Wells' novel broadcast in 1938 over the CBS Radio Network led to outrage because it supposedly suggested to many listeners that an actual alien invasion by Martians was currently in progress. In the wake of this, conspiracy theories on the presence of extraterrestrials became a widespread phenomenon in the United States during the 1940s and the beginning Space Age during the 1950s, accompanied by a surge of UFO reports. The term UFO itself was coined in 1952 in the context of the enormous popularity of the concept of "flying saucers" in the wake of the Kenneth Arnold UFO sighting in 1947. The Majestic 12 documents published in 1982 suggest that there was genuine interest in UFO conspiracy theories in the US government during the 1940s.

The trend to assume that celestial bodies were populated almost by default was tempered as actual probes visited potential alien abodes in the Solar System beginning in the second half of the 20th century, and by the 1970s belief in UFOs had become part of the fringe beliefs associated with the paranormal, New Age, Earth mysteries, Forteana etc. A number of UFO religions developed during the surge in UFO belief during th 1950s to 1970s period, and some, such as Scientology (founded 1953) and Raëlism (founded 1974) remain active into the present. The idea of "paleocontact", supposing that extraterrestrials ("ancient astronauts") have visited the Earth in the remote past and left traces in ancient cultures, appears in early 20th century fiction such as *The Call of Cthulhu* (1926), but it was given serious consideration in 1966 by astrophysicists I.S. Shklovski and Carl Sagan, and the idea came to be established as a notable aspect of the Ufology subculture in the wake of Erich von Däniken's *Chariots of the Gods?* (1968). Alien abduction claims were also widespread during the 1960s and 1970s in the United States.

On the scientific side, the possibility of extraterrestrial life on the Moon was decisively ruled out by the 1960s, and during the 1970s it became clear that most of the other bodies of the Solar System do not harbour highly developed life, although the question of primitive life on bodies in the Solar System remains an open question. Carl Sagan, Bruce Murray, and Louis Friedman founded the U.S. Planetary Society, partly as a vehicle for SETI studies in 1980, and since the 1990s, systematic search for radio signals attributable to intelligent extraterrestrial life has been ongoing.

Recent history

The failure of the SETI program to announce an intelligent radio signal after decades of effort has at least partially dimmed the prevailing optimism of the beginning of the space age. Notwithstanding, the unproven belief in extraterrestrial beings continues to be voiced in pseudoscience, conspiracy theories, and in popular folklore, notably "Area 51" and legends, but has also become a pop culture trope given less-than-serious treatment in popular entertainment with e.g. the ALF TV series (1986–1990), *The X-Files* (1993–2002), etc.

The SETI program is not the result of a continuous, dedicated search, but instead utilizes what resources and manpower it can, when it can. Furthermore, the SETI program only searches a limited range of frequencies at any one time.

In the words of SETI's Frank Drake, "All we know for sure is that the sky is not littered with powerful microwave transmitters". Drake has also noted that it is entirely possible that advanced technology results in communication being carried out in some way other than conventional radio transmission. At the same time, the data returned by space probes, and giant strides in detection methods, have allowed science to begin delineating habitability criteria on other worlds, and to confirm that at least other planets are plentiful, though aliens remain a question mark. The Wow! signal, from SETI, remains a speculative debate.

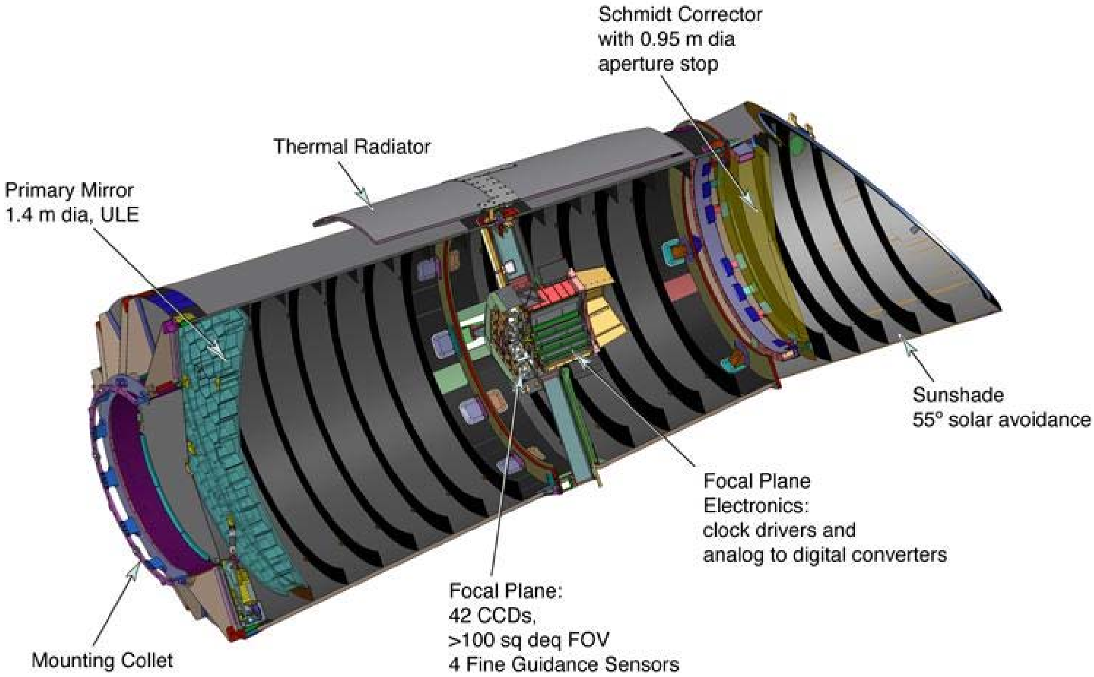
In 2000, geologist and paleontologist Peter Ward and astrobiologist Donald Brownlee published a book entitled *Rare Earth: Why Complex Life is Uncommon in the Universe*. In it, they discussed the Rare Earth hypothesis, in which they claim that Earth-like life is rare in the Universe, while microbial life is common. Ward and Brownlee are open to the idea of evolution on other planets which is not based on essential Earth-like characteristics (such as DNA and carbon).

The possible existence of primitive (microbial) life outside of Earth is much less controversial to mainstream scientists, although, at present, no direct evidence of such life has been found. Indirect evidence has been offered for the current existence of primitive life on Mars. However, the conclusions that should be drawn from such evidence remain in debate.

In September 2010, it was reported that the U.N. General Assembly had appointed Mazlan Othman as their official extraterrestrial liaison by the UK paper *The Sunday Times*. This claim was later refuted.

Theoretical physicist Stephen Hawking in 2010 warned that humans should not try to contact alien life forms. He warned that aliens might pillage Earth for resources. "If aliens visit us, the outcome would be much as when Columbus landed in America, which didn't turn out well for the Native Americans," he said. Geographer Jared Diamond has also expressed similar concerns.

Scientific search for extraterrestrial life



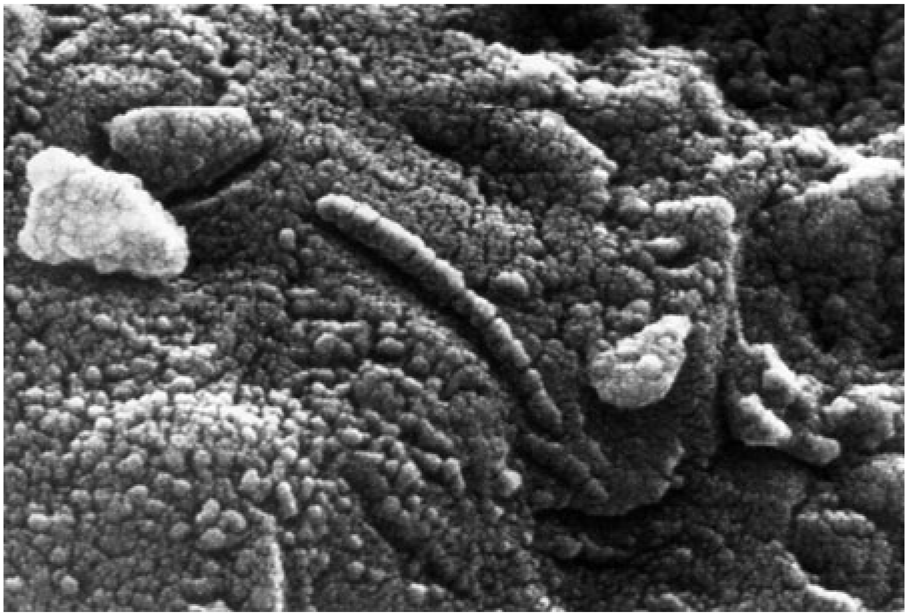
The NASA Kepler Mission for the search of extrasolar planets.

The scientific search for extraterrestrial life is being carried out both directly and indirectly.

Direct search

Scientists are directly searching for evidence of unicellular life within the Solar System, carrying out studies on the surface of Mars and examining meteors which have fallen to Earth. A mission is also proposed to Europa, one of Jupiter's moons with a possible liquid water layer under its surface, which might contain life.

There is some limited evidence that microbial life might possibly exist (or have existed) on Mars. An experiment on the Viking Mars lander reported gas emissions from heated Martian soil that some argue are consistent with the presence of microbes. However, the lack of corroborating evidence from other experiments on the Viking lander indicates that a non-biological reaction is a more likely hypothesis. Independently, in 1996, structures resembling nanobacteria were reportedly discovered in a meteorite, ALH84001, thought to be formed of rock ejected from Mars. This report is also controversial.



Electron micrograph of martian meteorite ALH84001 showing structures that some scientists think could be fossilized bacteria-like life forms.

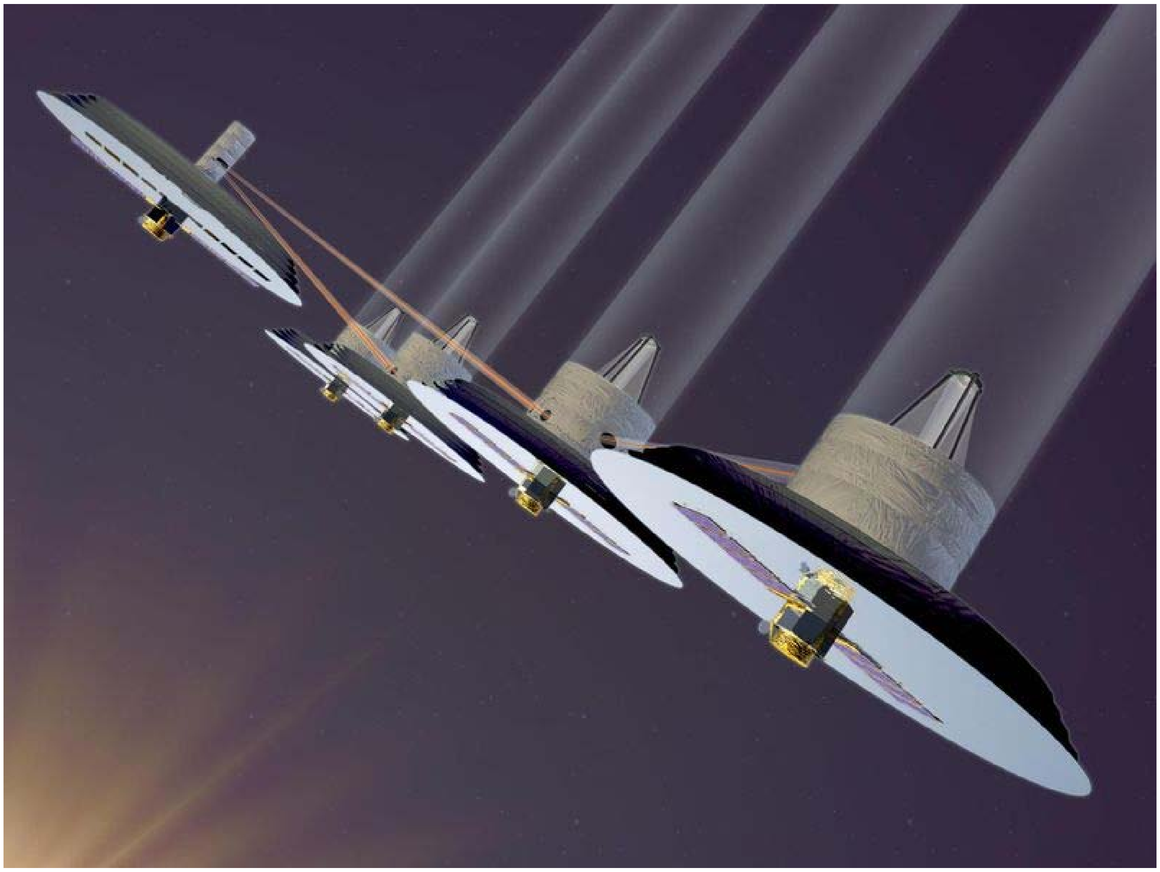
In February 2005, NASA scientists reported that they had found strong evidence of present life on Mars. The two scientists, Carol Stoker and Larry Lemke of NASA's Ames Research Center, based their claims on methane signatures found in Mars' atmosphere resembling the methane production of some forms of primitive life on Earth, as well as on their own study of primitive life near the Rio Tinto river in Spain. NASA officials soon denied the scientists' claims, and Stoker herself backed off from her initial assertions.

Though such findings are still very much in debate, support among scientists for the belief in the existence of life on Mars seems to be growing. In an informal survey conducted at the conference at which the European Space Agency presented its findings, 75 percent of the scientists in attendance were reported to believe that life once existed on Mars, and 25 percent reported a belief that life currently exists there.

The Gaia hypothesis stipulates that any planet with a robust population of life will have an atmosphere in chemical disequilibrium, which is relatively easy to determine from a distance by spectroscopy. However, significant advances in the ability to find and resolve light from smaller rocky worlds near their star are necessary before such spectroscopic methods can be used to analyze extrasolar planets.

On March 5, 2011, Richard B. Hoover, an astrobiologist with the Marshall Space Flight Center, speculated on the finding of alleged microfossils similar to cyanobacteria in CI1 carbonaceous meteorites. However, NASA formally distanced itself from Hoover's claim.

Indirect search



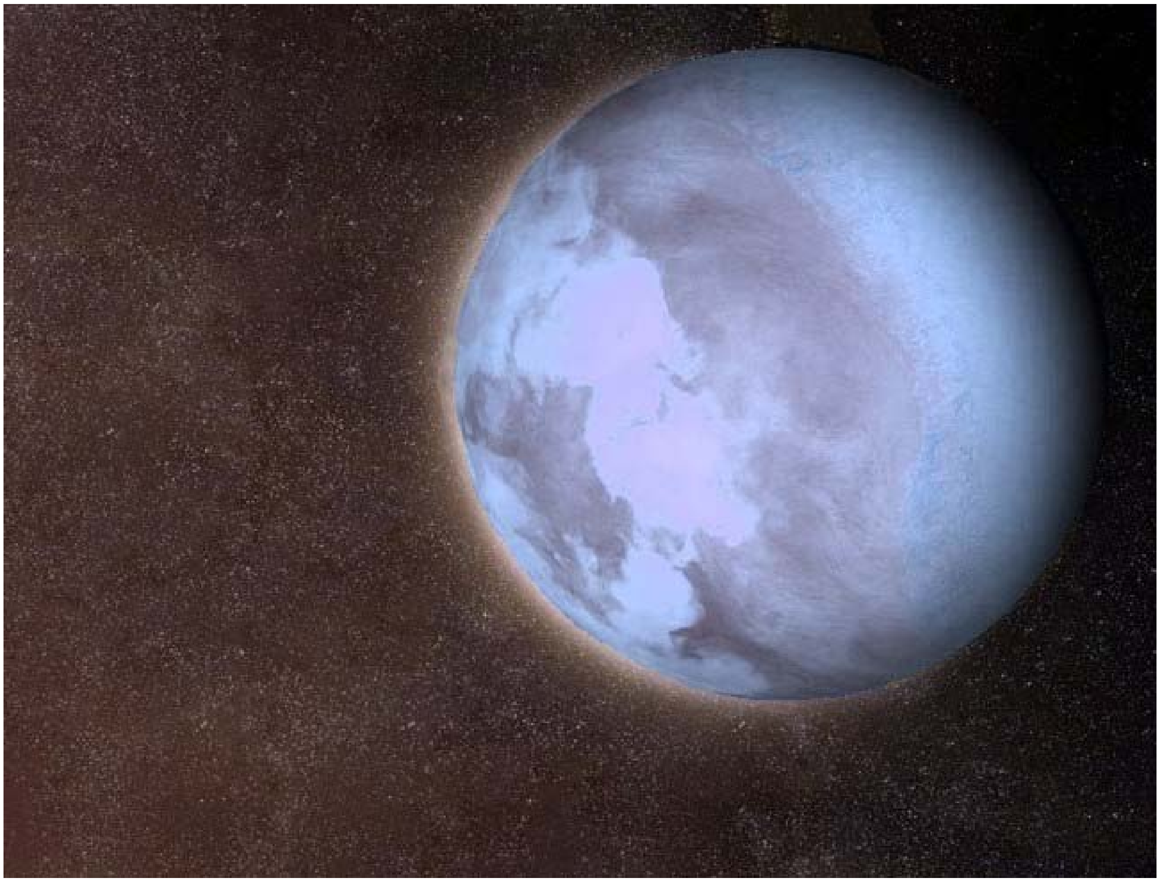
Terrestrial Planet Finder - A planned Infrared interferometer for finding Earth-like extrasolar planets (as of 2010, it has not received the funding from NASA which it needs—that funding is going towards the Kepler mission).

It is theorized that any technological society in space will be transmitting information. However, since there are no human systems, in general, intentionally transmitting information randomly into deep space, there is no guarantee that any other species would do so either. Also, the length of time required for a signal to travel across the vastness of space means that any signal detected, or not detected, would come from the distant past.

Nevertheless, projects such as SETI are conducting an astronomical search for radio activity which would confirm the presence of intelligent life. A related suggestion is that aliens might broadcast pulsed and continuous laser signals in the optical, as well as infrared, spectrum; laser signals have the advantage of not "smearing" in the interstellar medium, and may prove more conducive to communication between the stars. While other communication techniques, including laser transmission and interstellar spaceflight, have been discussed seriously and may well be feasible, the measure of effectiveness is the amount of information communicated per unit cost. This results in radio transmission as the method of choice.

Extrasolar planets

Astronomers also search for extrasolar planets that they believe would be conducive to life, such as Gliese 581 c, Gliese 581 g, Gliese 581 d and OGLE-2005-BLG-390Lb, which have been found to have Earth-like qualities. Current radiodetection methods have been inadequate for such a search, since the resolution afforded by recent technology is inadequate for a detailed study of extrasolar planetary objects. Future telescopes should be able to image planets around nearby stars, which may reveal the presence of life – either directly or through spectrography – and would reveal key information, such as the presence of free oxygen in a planet's atmosphere:



Artist's Impression of Gliese 581 c, the first extrasolar planet discovered within its star's habitable zone.

- Darwin was a proposed ESA mission designed to find Earth-like planets and analyze their atmosphere.
- The COROT mission, initiated by the French Space Agency, was launched in 2006, and is currently looking for extrasolar planets; it is the first of its kind.
- The Terrestrial Planet Finder was supposed to have been launched by NASA, but as of 2010, budget cuts have caused it to be delayed indefinitely.
- The Kepler Mission, largely replacing the Terrestrial Planet Finder, was launched in March 2009.

It has been argued that Alpha Centauri, the closest star system to Earth, may contain planets which could be capable of sustaining life.

On April 24, 2007, scientists at the European Southern Observatory in La Silla, Chile said they had found the first Earth-like planet. The planet, known as Gliese 581 c, orbits within the habitable zone of its star Gliese 581, a red dwarf star which is 20.5 light years (194 trillion km) from the Earth. It was initially thought that this planet could contain liquid water, but recent computer simulations of the climate on Gliese 581 c by Werner von Bloh and his team at Germany's Institute for Climate Impact Research suggest that carbon dioxide and methane in the atmosphere would create a runaway greenhouse effect. This would warm the planet well above the boiling point of water (100 degrees Celsius/212 degrees Fahrenheit), thus dimming the hopes of finding life. As a result of greenhouse models, scientists are now turning their attention to Gliese 581 d, which lies just outside of the star's traditional habitable zone.

On May 29, 2007, the Associated Press released a report stating that scientists identified twenty-eight new extra-solar planetary bodies. One of these newly-discovered planets is said to have many similarities to Neptune.

Extrasolar planets have been discovered on a regular basis since 1992, with 539 confirmed as of April 4, 2011.

The Drake equation

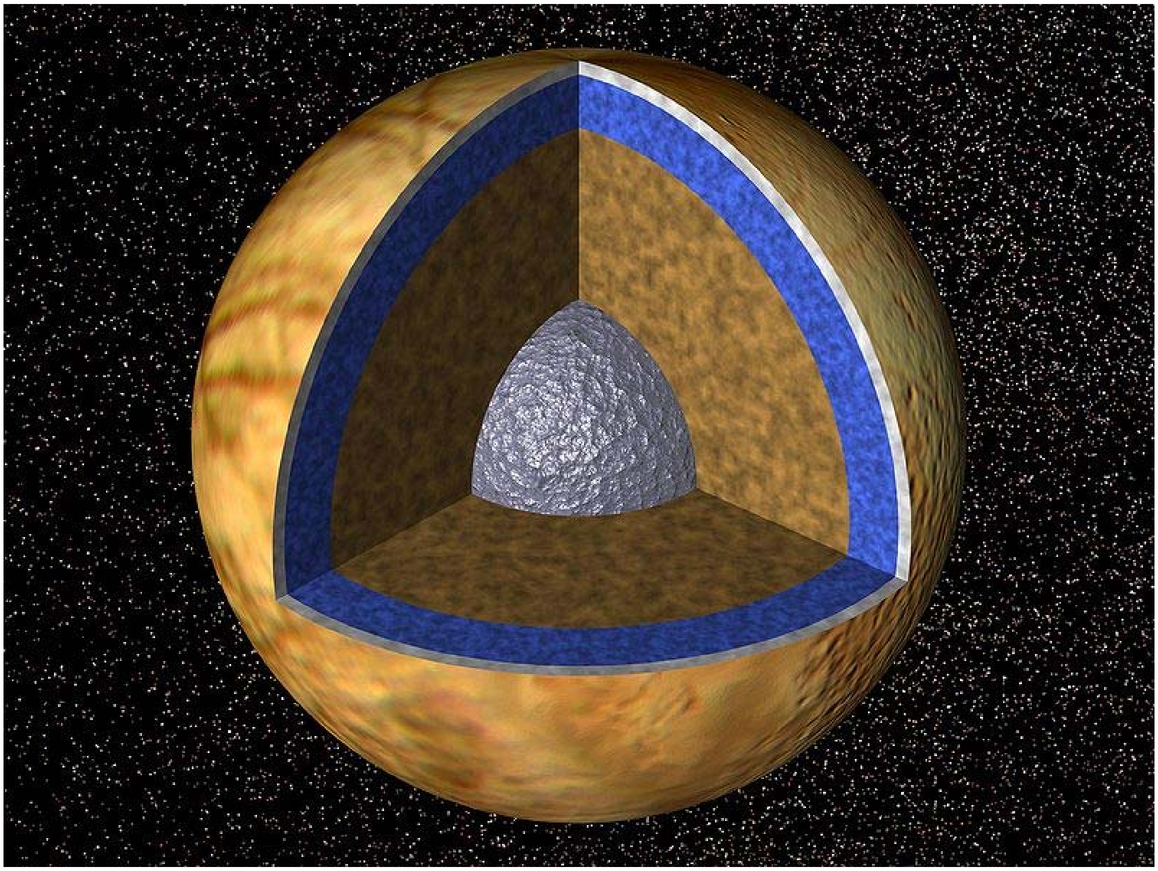
In 1961, University of California, Santa Cruz astronomer and astrophysicist Dr. Frank Drake devised the Drake equation. This controversial equation multiplied estimates of the following terms together:

- The rate of formation of suitable stars.
- The fraction of those stars which are orbited by planets.
- The number of Earth-like worlds per planetary system.
- The fraction of planets where intelligent life develops.
- The fraction of possible communicative planets.
- The "lifetime" of possible communicative civilizations.

Drake used the equation to estimate that there are approximately 10,000 planets in the Milky Way galaxy containing intelligent life with the possible capability of communicating with Earth.

Based on observations from the Hubble Space Telescope, there are at least 125 billion galaxies in the Universe. It is estimated that at least ten percent of all sun-like stars have a system of planets, i.e. there are 6.25×10^{18} stars with planets orbiting them in the Universe. Even if we assume that only one out of a billion of these stars have planets supporting life, there would be some 6.25×10^{9} (billion) life-supporting planetary systems in the Universe.

Possible extraterrestrial life in the Solar System



Europa, due to the possibility of an ocean under its icy crust, might host some form of microbial life.

Many bodies in the Solar System have been suggested as being capable of containing conventional organic life. The most commonly suggested ones are listed below; of these, five of the ten are moons, and are thought to have large bodies of underground liquid (streams), where life may have evolved in a way similar to biological communities surrounding deep sea vents.

- Mars — Life on Mars has been long speculated. Liquid water is widely thought to have existed on Mars in the past, and there may still be liquid water beneath the surface. Methane was found in the atmosphere of Mars and work is underway to discern its biological or abiotic origin. By July 2008, laboratory tests aboard NASA's Phoenix Mars Lander had identified water in a soil sample. The lander's robotic arm delivered the sample to an instrument which identifies vapors produced by the heating of samples. Recent photographs from the Mars Global Surveyor show evidence of recent (i.e. within 10 years) flows of a liquid on the Red Planet's frigid surface.
- Mercury — The MESSENGER expedition to Mercury has discovered that a large amount of water exists in its exosphere.

- Venus — Recently, scientists have speculated on the existence of microbes in the stable cloud layers 50 km above the surface, evidenced by hospitable climates and chemical disequilibrium.
- Jupiter — Carl Sagan and others in the 1960s and 70s computed conditions for hypothetical amino acid-based macroscopic life in the atmosphere of Jupiter, based on observed conditions of this atmosphere. These investigations inspired some science fiction stories.
- Ganymede (Largest moon of Jupiter) — Possible underground ocean.
- Callisto (Moon of Jupiter) — Possible underground ocean.
- Europa (Moon of Jupiter) — Europa may contain liquid water beneath its thick ice layer. It is possible that vents on the bottom of the ocean warm the ice, so liquid could exist beneath the ice layer, perhaps capable of supporting microbes and simple plants, just like in Earth's hydrothermal vents.
- Titan (Largest moon of Saturn) — The only known moon with a significant atmosphere. Data from the Cassini-Huygens mission refuted the hypothesis of a global hydrocarbon ocean, but later demonstrated the existence of liquid hydrocarbon lakes in the polar regions—the first liquid lakes discovered outside of Earth. Analysis of data from the mission has uncovered aspects of atmospheric chemistry near the surface which are consistent with —but do not prove— the hypothesis that organisms there are consuming hydrogen, acetylene and ethane, and producing methane.
- Enceladus (Moon of Saturn) — Geothermal activity, water vapor. Possible under-ice oceans heated by tidal effects.

Numerous other bodies have been suggested as potential hosts for microbial life. Fred Hoyle has proposed that life might exist on comets, as some Earth microbes managed to survive on a lunar probe for many years (later considered doubtful as sterile procedures may not have been fully followed). However, it is considered highly unlikely that complex multicellular organisms of the conventional chemistry of terrestrial life (i.e. animals and plants) could exist under these living conditions.

Even if microbial extraterrestrial life were found on another body in the Solar System, it would still need to be proven that such life did not originate from Earth in the recent or distant past. For example, an alternate explanation for the hypothetical existence of microbial life on Titan has already been formally proposed—theorizing that microorganisms could have left Earth when it suffered a massive asteroid or comet impact (such as the impact that created Chicxulub crater only 65 mya), and survived a journey through space to land on Titan 1 million years later. The Living Interplanetary Flight Experiment, developed by the Planetary Society and due to be launched in 2011, has been designed to test similar theories.

Chapter 4

Cambrian Explosion

The **Cambrian explosion** or **Cambrian radiation** was the relatively rapid appearance, over a period of many million years, of most major Phyla around 530 million years ago, as found in the fossil record. This was accompanied by a major diversification of other organisms, including animals, phytoplankton, and calcimicrobes. Before about 580 million years ago, most organisms were simple, composed of individual cells occasionally organized into colonies. Over the following 70 or 80 million years the rate of evolution accelerated by an order of magnitude (as defined in terms of the extinction and origination rate of species) and the diversity of life began to resemble today's.

The Cambrian explosion has generated extensive scientific debate. The seemingly rapid appearance of fossils in the "Primordial Strata" was noted as early as the mid 19th century, and Charles Darwin saw it as one of the main objections that could be made against his theory of evolution by natural selection.

The long-running puzzlement about the appearance of the Cambrian fauna, seemingly abruptly and from nowhere, centers on three key points: whether there really was a mass diversification of complex organisms over a relatively short period of time during the early Cambrian; what might have caused such rapid change; and what it would imply about the origin and evolution of animals. Interpretation is difficult due to a limited supply of evidence, based mainly on an incomplete fossil record and chemical signatures left in Cambrian rocks.

History and significance

Geologists as long ago as William Buckland (1784–1856) realised that a dramatic step-change in the fossil record occurred around the base of what we now call the Cambrian. Charles Darwin considered this sudden appearance of many animal groups with few or no antecedents to be the greatest single objection to his theory of evolution. He had even devoted a substantial chapter of *The Origin of Species* to solving this problem.

American palæontologist Charles Walcott proposed that an interval of time, the “Lipalian”, was not represented in the fossil record or did not preserve fossils, and that the ancestors of the Cambrian animals evolved during this time.

More recently it was discovered that the history of life on earth goes back at least 3,450 million years: rocks of that age at Warrawoona in Australia contain fossils of stromatolites, stubby pillars that are formed by colonies of micro-organisms. Fossils (*Grypania*) of more complex eukaryotic cells, from which all animals, plants and fungi are built, have been found in rocks from 1,400 million years ago, in China and Montana. Rocks dating from 565 to 543 million years ago contain fossils of the Ediacara biota, organisms so large that they must have been multi-celled, but very unlike any modern organism. P. E. Cloud argued in 1948 that there was a period of "eruptive" evolution in the Early Cambrian, but as recently as the 1970s there was no sign of how the *relatively* modern-looking organisms of the Middle and Late Cambrian arose.

Opabinia made the largest single contribution to modern interest in the Cambrian explosion.

The intense modern interest in this "Cambrian explosion" was sparked by the work of Harry B. Whittington and colleagues, who in the 1970s re-analysed many fossils from the Burgess Shale and concluded that several were complex but different from any living animals. The most common organism, *Marrella*, was clearly an arthropod, but not a member of any known arthropod class. Organisms such as the five-eyed *Opabinia* and spiny slug-like *Wiwaxia* were so different from anything else known that Whittington's team assumed they must represent different phyla, only distantly related to anything known today. Stephen Jay Gould's popular 1989 account of this work, *Wonderful Life*, brought the matter into the public eye and raised questions about what the explosion represented. While differing significantly in details, both Whittington and Gould proposed that all modern animal phyla had appeared rather suddenly. This view was influenced by the theory of punctuated equilibrium, which Eldredge and Gould developed in the early 1970s and which views evolution as long intervals of near-stasis "punctuated" by short periods of rapid change.

Other analyses, some more recent and some dating back to the 1970s, argue that complex animals similar to modern types evolved well before the start of the Cambrian. There has also been intense debate whether there was a genuine "explosion" of modern forms in the Cambrian and, to the extent that there was, how it happened and why it happened then.

Types of evidence

Deducing the events of half a billion years ago is difficult, as evidence comes exclusively from biological and chemical signatures in rocks and very sparse fossils.

Dating the Cambrian

Accurate absolute radiometric dates for much of the Cambrian, obtained by detailed analysis of radioactive elements contained within rocks, have only rather recently become available, and for only a few regions.

Relative dating (*A* was before *B*) is often sufficient for studying processes of evolution, but this too has been difficult, because of the problems involved in matching up rocks of the same age across different continents.

Therefore dates or descriptions of sequences of events should be regarded with some caution until better data becomes available.

Body fossils

Fossils of organisms' bodies are usually the most informative type of evidence. Fossilisation is a rare event, and most fossils are destroyed by erosion or metamorphism before they can be observed. Hence the fossil record is very incomplete, increasingly so, further back in time. Despite this, they are often adequate to illustrate the broader patterns of life's history. There are also biases in the fossil record: different environments are more favourable to the preservation of different types of organism or parts of organisms. Further, only the parts of organisms that were already mineralised are usually preserved, such as the shells of molluscs. Since most animal species are soft-bodied, they decay before they can become fossilised. As a result, although there are 30-plus phyla of living animals, two-thirds have never been found as fossils.

This *Marrella* specimen illustrates how clear and detailed the fossils from the Burgess Shale lagerstätte are.

The Cambrian fossil record includes an unusually high number of lagerstätten, which preserve soft tissues. These allow palæontologists to examine the internal anatomy of animals which in other sediments are only represented by shells, spines, claws, etc. – if they are preserved at all. The most significant Cambrian lagerstätten are the early Cambrian Maotianshan shale beds of Chengjiang (Yunnan, China) and Sirius Passet (Greenland); the middle Cambrian Burgess Shale (British Columbia, Canada); and the late Cambrian Orsten (Sweden) fossil beds.

While lagerstätten preserve far more than the conventional fossil record, they are far from complete. Because lagerstätten are restricted to a narrow range of environments (where soft-bodied organisms can be preserved very quickly, e.g. by mudslides), most animals are probably not represented; further, the exceptional conditions that create lagerstätten probably do not represent normal living conditions. In addition, the known Cambrian lagerstätten are rare and difficult to date, while Precambrian lagerstätten have yet to be studied in detail.

The sparseness of the fossil record means that organisms usually exist long before they are found in the fossil record – this is known as the Signor-Lipps effect.

Trace fossils

Rusophycus and other trace fossils from the Gog Formation, Middle Cambrian, Lake Louise, Alberta, Canada.

Trace fossils consist mainly of tracks and burrows, but also include coprolites (fossil feces) and marks left by feeding. Trace fossils are particularly significant because they represent a data source that is not limited to animals with easily-fossilized hard parts, and which reflects organisms' behaviour. Also many traces date from significantly earlier than the body fossils of animals that are thought to have been capable of making them. Whilst exact assignment of trace fossils to their makers is generally impossible, traces may for example provide the earliest physical evidence of the appearance of moderately complex animals (comparable to earthworms).

Geochemical observations

Several chemical markers indicate a drastic change in the environment around the start of the Cambrian. The markers are consistent with a mass extinction, or with a massive warming resulting from the release of methane ice. Such changes may reflect a cause of the Cambrian explosion, although they may also have resulted from an increased level of biological activity – a possible result of the explosion. Despite these uncertainties, the geochemical evidence helps by making scientists focus on theories that are consistent with at least one of the likely environmental changes.

Phylogenetic techniques

Cladistics is a technique for working out the “family tree” of a set of organisms. It works by the logic that, if groups B and C have more similarities to each other than either has to group A, then B and C are more closely related to each other than either is to A. Characteristics which are compared may be anatomical, such as the presence of a notochord, or molecular, by comparing sequences of DNA or protein. The result of a successful analysis is a hierarchy of clades – groups whose members are believed to share a common ancestor. The cladistic technique is sometimes fallible, as some features, such as wings or camera eyes, evolved more than once, convergently – this must be taken into account in analyses.

From the relationships, it may be possible to constrain the date that lineages first appeared. For instance, if fossils of B or C date to X million years ago and the calculated "family tree" says A was an ancestor of B and C, then A must have evolved more than X million years ago.

It is also possible to estimate how long ago two living clades diverged – i.e. approximately how long ago their last common ancestor must have lived – by assuming that DNA mutations accumulate at a constant rate. These "molecular clocks", however, are fallible, and provide only a very approximate timing: they are not sufficiently precise and reliable for estimating when the groups that feature in the Cambrian explosion first evolved, and estimates produced by different techniques vary by a factor of two. However, the clocks can give an indication of branching rate, and when combined with the constraints of the fossil record, recent clocks suggest a sustained period of diversification through the Ediacaran and Cambrian.

Explanation of a few scientific terms

A **phylum** is the highest level in the Linnean system for classifying organisms. Phyla can be thought of as groupings of animals based on general body plan. Despite the seemingly different *external* appearances of organisms, they are classified into phyla based on their *internal* and developmental organizations. For example, despite their obvious differences, spiders and barnacles both belong to the phylum Arthropoda; but earthworms and tapeworms, although similar in shape, belong to different phyla.

A phylum is not a fundamental division of nature, such as the difference between electrons and protons. It is simply a very high-level grouping in a classification system created to describe all currently living organisms. This system is imperfect, even for modern animals: different books quote different numbers of phyla, mainly because they disagree about the classification of a huge number of worm-like species. As it is based on living organisms, it accommodates extinct organisms poorly, if at all.

The concept of **stem groups** was introduced to cover evolutionary "aunts" and "cousins" of living groups. A crown group is a group of closely-related living animals plus their last common ancestor plus all its descendants. A stem group is a set of offshoots from the lineage at a point earlier than the last common ancestor of the crown group; it is a *relative* concept, for example tardigrades are living animals which form a crown group in their own right, but Budd (1996) regarded them also as being a stem group relative to the arthropods.

Triploblastic means consisting of 3 layers, which are formed in the embryo, quite early in the animal's development from a single-celled egg to a larva or juvenile form. The innermost layer forms the digestive tract (gut); the outermost forms skin; and the middle one forms muscles and all the internal organs except the digestive system. Most types of living animal are triploblastic – the best-known exceptions are Porifera (sponges) and Cnidaria (jellyfish, sea anemones, etc.).

The **bilaterians** are animals which have right and left sides at some point in their life history. This implies that they have top and bottom surfaces and, importantly, distinct front and back ends. All known bilaterian animals are triploblastic, and all known triploblastic animals are bilaterian. Living Echinoderms (sea stars, sea urchins, sea cucumbers, etc.) *look* radially symmetrical (like wheels) rather than bilaterian, but their larvae exhibit bilateral symmetry and some of the earliest echinoderms may have been bilaterally symmetrical. Porifera and Cnidaria are radially symmetrical, non-bilaterian and non-triploblastic.

Coelomate means having a body cavity (coelom) which contains the internal organs. Most of the phyla featured in the debate about the Cambrian explosion are coelomates: arthropods, annelid worms, molluscs, echinoderms and chordates – the non-coelomate priapulids are an important exception. All known coelomate animals are triploblastic bilaterians, but some triploblastic bilaterian animals do not have a coelom – for example flatworms, whose organs are surrounded by unspecialized tissues.

Precambrian life

Our understanding of the Cambrian explosion relies upon knowing what was there beforehand – did the event herald the sudden appearance of a wide range of animals and behaviours, or did such things exist beforehand?

Evidence of animals around 1 billion years ago

Stromatolites (Pika Formation, Middle Cambrian) near Helen Lake, Banff National Park, Canada.

Modern stromatolites in Hamelin Pool Marine Nature Reserve, Western Australia.

Changes in the abundance and diversity of some types of fossil have been interpreted as evidence for "attacks" by animals or other organisms. Stromatolites, stubby pillars built by colonies of microorganisms, are a major constituent of the fossil record from about 2,700 million years ago, but their abundance and diversity declined steeply after about 1,250 million years ago. This decline has been attributed to disruption by grazing and burrowing animals.

Precambrian marine diversity was dominated by small fossils known as acritarchs. This term describes almost any small organic walled fossil–from the egg cases of small metazoans to resting cysts of many different kinds of green algae. After appearing around 2,000 million years ago, acritarchs underwent a boom around 1,000 million years ago, increasing in abundance, diversity, size, complexity of shape and especially size and number of spines. Their increasingly spiny forms in the last 1 billion years may indicate an increased need for defence against predation. Other groups of small organisms from the Neoproterozoic era also show signs of anti-predator defenses. A consideration of taxon longevity appears to support an increase in predation pressure around this time, However, in general, the rate of evolution in the Precambrian was very slow, with many cyanobacterial species persisting unchanged for billions of years.

If these predatory organisms really were metazoans, this means that Cambrian animals did not appear "from no-where" at the base of the Cambrian; their predecessors had existed for hundreds of millions of years.

Fossils of the Doushantuo formation

The 580 million year old Doushantuo formation harbours microscopic fossils which may represent early bilaterians. Some have been described as animal embryos and eggs, although some of these may represent the remains of giant bacteria. Another fossil, *Vernanimalcula*, has been interpreted as a coelomate bilaterian, but may simply be an infilled bubble.

These fossils form the earliest hard-and-fast evidence of animals, as opposed to other predators.

Burrows

An Ediacaran trace fossil, made when an organism burrowed below a microbial mat.

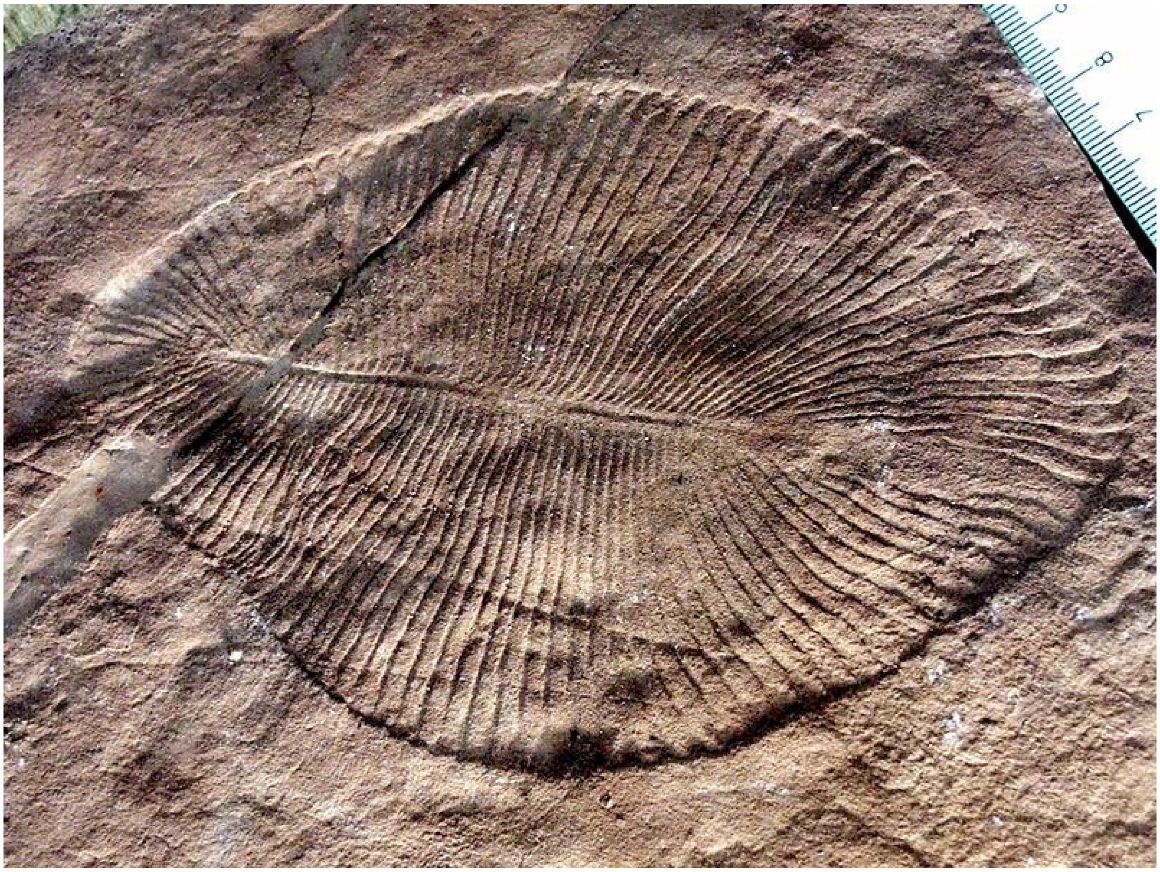
The traces of organisms moving on and directly underneath the microbial mats that covered the Ediacaran sea floor are preserved from the Ediacaran period, about 565 million years ago. They were probably made by organisms resembling earthworms in shape, size, and how they moved. The burrow-makers have never been found preserved, but because they would need a head and a tail, the burrowers probably had bilateral symmetry – which would in all probability make them bilaterian animals. They fed above the sediment surface, but were forced to burrow to avoid predators.

Around the start of the Cambrian (about 542 million years ago) many new types of traces first appear, including well-known vertical burrows such as *Diplocraterion* and *Skolithos*, and traces normally attributed to arthropods, such as *Cruziana* and *Rusophycus*. The vertical burrows indicate that worm-like animals acquired new behaviours, and possibly new physical capabilities. Some Cambrian trace fossils indicate that their makers possessed hard exoskeletons, although they were not necessarily mineralised.

Burrows provide firm evidence of complex organisms; they are also much more readily preserved than body fossils, to the extent that the absence of trace fossils has been used to imply the genuine absence of large, motile bottom-dwelling organisms. They provide a further line of evidence to show that the Cambrian explosion represents a real diversification, and is not a preservational artefact.

Indeed, as burrowing became established, it allowed an explosion of its own, for as burrowers disturbed the sea floor, they aerated it, mixing oxygen into the toxic muds. This made the bottom sediments more hospitable, and allowed a wider range of organisms to inhabit them – creating new niches and the scope for higher diversity.

Ediacaran organisms



Dickinsonia costata, an Ediacaran organism of unknown affinity, with a quilted appearance.

At the start of the Ediacaran period, much of the acritarch fauna, which had remained relatively unchanged for hundreds of millions of years, became extinct, to be replaced with a range of new, larger species which would prove far more ephemeral. This radiation, the first in the fossil record, is followed soon after by an array of unfamiliar, large, fossils dubbed the Ediacara biota, which flourished for 40 million years until the start of the Cambrian. Most of this "Ediacara biota" were at least a few centimeters long, significantly larger than any earlier fossils. The organisms form three distinct assemblages, increasing in size and complexity as time progresses.

Many of these organisms were quite unlike anything that appeared before or since, resembling discs, mud-filled bags, or quilted mattresses – one palæontologist proposed that the strangest organisms should be classified as a separate kingdom, Vendozoa.

Fossil of *Kimberella*, a triploblastic bilaterian , and possibly a mollusc.

At least some may have been early forms of the phyla at the heart of the "Cambrian explosion" debate, having been interpreted as early molluscs (*Kimberella*), echinoderms (*Arkarua*); and arthropods (*Spriggina*, *Parvancorina*). There is still debate about the classification of these specimens, mainly because the diagnostic features which allow taxonomists to classify more recent organisms, such as similarities to living organisms, are generally absent in the Ediacarans. However there seems little doubt that *Kimberella* was at least a triploblastic bilaterian animal. These organisms are central to the debate about how abrupt the Cambrian explosion was. If some were early members of the animal phyla seen today, the "explosion" looks a lot less sudden than if all these organisms represent an unrelated "experiment", and were replaced by the animal kingdom fairly soon thereafter (40M years is "soon" by evolutionary and geological standards).

Ediacaran–Early Cambrian skeletalization

The first Ediacaran and lowest Cambrian (Nemakit-Daldynian) skeletal fossils represent tubes and problematic sponge spicules. The oldest sponge spicules are monaxon siliceous, aged around 580 million years ago, known from the Doushantou Formation in the China and from deposits of the same age in Mongolia. In the late Ediacaran-lowest Cambrian, numerous tube dwellings of enigmatic organisms appeared. It was organic-walled tubes (e.g. *Saarina*) and chitinous tubes of the sabelliditids (e.g. *Sokoloviina*, *Sabellidites*, *Paleolina*) which prospered up to the beginning of the Tommotian. The mineralized tubes of *Cloudina*, *Namacalathus*, *Sinotubulites* and a dozen more of the other organisms from carbonate rocks formed near the end of the Ediacaran period from 549 to 542 million years ago, as well as the triradially symmetrical mineralized tubes of anabaritids (e.g. *Anabarites*, *Cambrotubulus*) from uppermost Ediacaran and lower Cambrian. It is interesting to notice that Ediacaran mineralized tubes are often found in carbonates of the stromatolite reefs and thrombolites, i.e. they could live in environment adverse for the majority of animals.

Although they are as hard to classify as most other Ediacaran organisms, they are important in two other ways. First, they are the earliest known calcifying organisms (organisms that built shells out of calcium carbonate). Secondly, these tubes are a device to rise over a substrate and competitors for effective feeding and, to a lesser degree, they serve as armor for protection against predators and adverse conditions of environment. Some *Cloudina* fossils show small holes in shells. It is possible that the holes are evidence of boring by predators sufficiently advanced to penetrate shells. A possible "evolutionary arms race" between predators and prey is one of the theories that attempt to explain the Cambrian explosion.

In the lowest Cambrian occurs the extinction of stromatolites. This allowed animals to begin colonization of warm-water pools with carbonate sedimentation. At first it was anabaritids and *Protohertzina* (the fossilized grasping spines of chaetognaths) fossils. Such mineral skeletons as shells, sclerites, thorns and plates appeared in uppermost Nemakit-Daldynian; they were the earliest species of halkierids, gastropods, hyoliths and other rare organisms. The beginning of the Tommotian has historically been understood to mark an explosive increase of the number and variety of fossils of mollusks, hyoliths and sponges, along with a rich complex of skeletal elements of unknown animals, the first archaeocyathids, brachiopods, tommotiids and others. This sudden increase is partially an artefact of missing strata at the Tommotian type section, and most of this fauna in fact began to diversify in a series of pulses through the Nemakit-Daldynian and into the Tommotian.

Some animals may already have had sclerites, thorns and plates in the Ediacaran (e.g. *Kimberella* had hard sclerites, probably of carbonate), but thin carbonate skeletons cannot be fossilized in siliciclastic deposits.

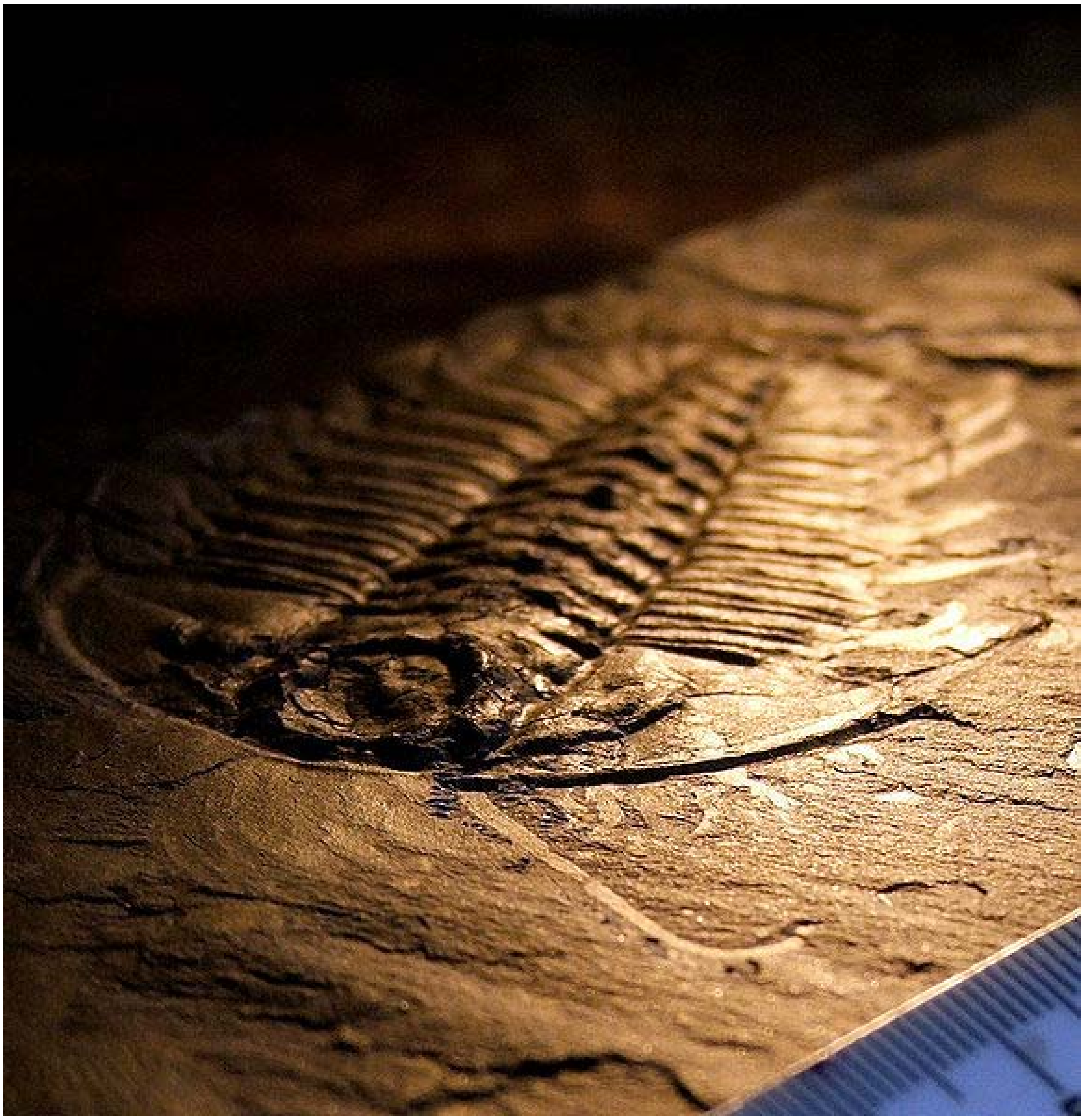
Cambrian life

Small shelly fauna

Fossils known as “small shelly fauna” have been found in many parts on the world, and date from just before the Cambrian to about 10 million years after the start of the Cambrian. These are a very mixed collection of fossils: spines, sclerites (armor plates), tubes, archeocyathids (sponge-like animals) and small shells very like those of brachiopods and snail-like molluscs – but all tiny, mostly 1 to 2 mm long.

While small, these fossils are far more common than complete fossils of the organisms that produced them; crucially, they cover the window from the start of the Cambrian to the first lagerstatten: a period of time that is otherwise lacking in fossils. Hence they supplement the conventional fossil record, and allow the fossil ranges of many groups to be extended.

Early Cambrian trilobites and echinoderms



A fossilized trilobite, an ancient type of arthropod. This specimen, from the Burgess shale, preserves "soft parts" – the antennae and legs.

The earliest Cambrian trilobite fossils are about 530 million years old, but the class was already quite diverse and worldwide, suggesting that they had been around for quite some time. It is important to remember that the fossil record of trilobites begins from the time of appearance of trilobites with mineral exoskeleton - not from the time of their origin.

The earliest generally-accepted echinoderm fossils appeared a little bit later, in the Late Atdabanian; unlike modern echinoderms, these early Cambrian echinoderms were not all radially symmetrical.

These provide firm data points for the "end" of the explosion, or at least indications that the crown groups of modern phyla were represented.

Burgess shale type faunas

The Burgess shale and similar lagerstatten preserve the soft parts of organisms, which provides a wealth of data to aid in the classification of enigmatic fossils. It often preserved complete specimens of organisms only otherwise known from dispersed parts, such as loose scales or isolated mouthparts. Further, the majority of organisms and taxa in these horizons are entirely soft bodied – hence absent from the rest of the fossil record. Since a large part of the ecosystem is preserved, the ecology of the community can also be tentatively reconstructed. However, the assemblages may represent a "museum": a deep water ecosystem that is evolutionarily "behind" the rapidly diversifying faunas of shallower waters.

Because the lagerstatten provide a mode and quality of preservation that's virtually absent outside of the Cambrian, lots of organisms appear completely different to anything known from the conventional fossil record. This led early workers in the field to attempt to shoehorn the organisms into extant phyla; the shortcomings of this approach led them to erect a multitude of new phyla to accommodate all the oddballs. It has since been realised that most oddballs diverged from lineages before they established the phyla we know today – slightly different designs, which were fated to perish rather than flourish into phyla, as their cousin lineages did.

The preservational mode is rare in the preceding Ediacaran period, but those assemblages known show no trace of animal life – perhaps implying a genuine absence of macroscopic metazoans.

Early Cambrian crustaceans

Crustaceans, one of the four great modern groups of arthropods, are very rare throughout the Cambrian. Convincing crustaceans were once thought to be common in Burgess shale-type biotas, but none of these individuals can be shown to fall into the crown group of "true crustaceans". The Cambrian record of crown group crustaceans comes from microfossils. The Swedish Orsten horizons contain later Cambrian crustacea, but only organisms smaller than 2 mm are preserved. This restricts the data set to juveniles and miniaturised adults.

A more informative data source is the organic microfossils of the Mount Cap formation, Canada. This late Early Cambrian assemblage (510 to 515 million years ago) consists of microscopic fragments of arthropods' cuticle, which is left behind when the rock is dissolved with a strong acid. The diversity of this assemblage is similar to that of modern crustacean faunas. Most interestingly, analysis of fragments of feeding machinery found in the formation shows that it was adapted to feed in a very precise and refined fashion. This contrasts with most other early Cambrian arthropods, which fed messily by shovelling anything they could get their feeding appendages on into their mouths. This sophisticated and specialised feeding machinery belonged to a large (~30 cm) organism, and would have provided great potential for diversification: specialised feeding apparatus allows a number of different approaches to feeding and development, and creates a number of different approaches to avoid being eaten

Early Ordovician radiation

After a mass extinction at the Cambrian-Ordovician boundary, another radiation occurred, which established the taxa which would dominate the Palaeozoic.

During this radiation, the total number of orders doubled, and families tripled, increasing marine diversity to levels typical of the Palaeozoic, and disparity to levels approximately equivalent to today's.

How real was the explosion?

The fossil record as Darwin knew it seemed to suggest that the major metazoan groups appeared in a few million years of the early to mid-Cambrian, and even in the 1980s this still appeared to be the case.

However, evidence of Precambrian metazoa is gradually accumulating. If the Ediacaran *Kimberella* was a mollusc-like protostome (one of the two main groups of coelomates), the protostome and deuterostome lineages must have split significantly before 550 million years ago (deuterostomes are the other main group of coelomates). Even if it is not a protostome, it is widely accepted as a bilaterian. Since fossils of rather modern-looking Cnidarians (jellyfish-like organisms) have been found in the Doushantuo lagerstätte, the Cnidarian and bilaterian lineages must have diverged well over 580 million years ago.

Trace fossils and predatory borings in *Cloudina* shells provide further evidence of Ediacaran animals. Some fossils from the Doushantuo formation have been interpreted as embryos and one (*Vernanimalcula*) as a bilaterian coelomate, although these interpretations are not universally accepted. Earlier still, predatory pressure has acted on stromatolites and acritarchs since around 1,250 million years ago.

The presence of Precambrian animals somewhat dampens the "bang" of the explosion: not only was the appearance of animals gradual, but their evolutionary radiation ("diversification") may also not have been as rapid as once thought. Indeed, statistical analysis shows that the Cambrian explosion was no faster than any of the other radiations in animals' history. However, it does seem that some innovations linked to the explosion — such as resistant armour — only evolved once in the animal lineage; this makes a lengthy Precambrian animal lineage harder to defend. Further, the conventional view that all the phyla arose in the Cambrian is flawed; whilst the phyla may have diversified in this time period, representatives of the crown-groups of many phyla do not appear until much later in the Phanerozoic. Further, the mineralized phyla that form the basis of the fossil record may not be representative of other phyla, since most mineralized phyla originated in a benthic setting. The fossil record is consistent with a Cambrian Explosion that was limited to the benthos, with pelagic phyla evolving much later.

There is little doubt that disparity – that is, the range of different organism "designs" or "ways of life" – rose sharply in the early Cambrian. However, recent research has overthrown the once-popular idea that disparity was exceptionally high throughout the Cambrian, before subsequently decreasing. In fact, disparity remains relatively low throughout the Cambrian, with modern levels of disparity only attained after the early Ordovician radiation.

The diversity of many Cambrian assemblages is similar to today's.

Possible causes of the “explosion”

Despite the evidence that moderately complex animals (triploblastic bilaterians) existed before and possibly long before the start of the Cambrian, it seems that the pace of evolution was exceptionally fast in the early Cambrian. Possible explanations for this fall into three broad categories: environmental, developmental, and ecological changes. Any explanation must explain the timing and magnitude of the explosion.

Changes in the environment

Increase in oxygen levels

Earth’s earliest atmosphere contained no free oxygen; the oxygen that animals breathe today, both in the air and dissolved in water, is the product of billions of years of photosynthesis. As a general trend, the concentration of oxygen in the atmosphere has risen gradually over about the last 2.5 billion years.

Shortage of oxygen might well have prevented the rise of large, complex animals. The amount of oxygen an animal can absorb is largely determined by the area of its oxygen-absorbing surfaces (lungs and gills in the most complex animals; the skin in less complex ones); but the amount needed is determined by its volume, which grows faster than the oxygen-absorbing area if an animal’s size increases equally in all directions. An increase in the concentration of oxygen in air or water would increase the size to which an organism could grow without its tissues becoming starved of oxygen. However, members of the Ediacara biota reached metres in length; clearly oxygen did not limit *their* growth. Other metabolic functions may have been inhibited by lack of oxygen, for example the construction of tissue such as collagen, required for the construction of complex structures, or to form molecules for the construction of a hard exoskeleton. However, animals are not affected when similar oceanographic conditions occur in the Phanerozoic; there is no convincing correlation between oxygen levels and evolution, so oxygen may have been no more a prerequisite to complex life than liquid water or primary productivity.

Snowball Earths

In the late Neoproterozoic (extending into the early Ediacaran period), the Earth suffered massive glaciations in which most of its surface was covered by ice. This may have caused a mass extinction, creating a genetic bottleneck; the resulting diversification may have given rise to the Ediacara biota, which appears soon after the last "Snowball Earth" episode. However, the snowball episodes occurred a long time before the start of the Cambrian, and it is hard to see how so much diversity could have been caused by even a series of bottlenecks; the cold periods may even have *delayed* the evolution of large size.

Developmental explanations

A range of theories are based on the concept that minor modifications to animals' development as they grow from embryo to adult may have been able to cause very large changes in the final adult form. The hox genes, for example, control which organs individual regions of an embryo will develop into. For instance, if a certain hox gene is expressed, a region will develop into a limb; if a different hox gene is expressed in that region (a minor change), it could develop into an eye instead (a phenotypically major change).

Such a system allows a large range of disparity to appear from a limited set of genes, but such theories linking this with the explosion struggle to explain why the origin of such a development system should by itself lead to increased diversity or disparity. Evidence of Precambrian metazoans combines with molecular data to show that much of the genetic architecture that could feasibly have played a role in the explosion was already well established by the Cambrian.

This apparent paradox is addressed in a theory that focuses on the physics of development. It is proposed that the emergence of simple multicellular forms provided a changed context and spatial scale in which novel physical processes and effects were mobilized by the products of genes that had previously evolved to serve unicellular functions. Morphological complexity (layers, segments, lumens, appendages) arose, in this view, by self-organization.

Ecological explanations

These focus on the interactions between different types of organism. Some of these hypotheses deal with changes in the food chain; some suggest arms races between predators and prey, and others focus on the more general mechanisms of coevolution. Such theories are well suited to explaining why there was a rapid increase in both disparity and diversity, but they must explain why the "explosion" happened when it did.

End-Ediacaran mass extinction

Evidence for such an extinction includes the disappearance from the fossil record of the Ediacara biota and shelly fossils such as *Cloudina*, and the accompanying perturbation in the $\delta^{13}C$ record. Mass extinctions are often followed by adaptive radiations as existing clades expand to occupy the ecospace emptied by the extinction. However, once the dust had settled, overall disparity and diversity returned to the pre-extinction level in each of the Phanerozoic extinctions.

Evolution of eyes

Andrew Parker has proposed that predator-prey relationships changed dramatically after eyesight evolved. Prior to that time hunting and evading were both close-range affairs – smell, vibration, and touch were the only senses used. When predators could see their prey from a distance, new defensive strategies were needed. Armor, spines, and similar defenses may also have evolved in response to vision. He further observed that where animals lose vision in unlighted environments such as caves, diversity of animal forms tends to decrease. Nevertheless many scientists doubt that vision could have caused the explosion. Eyes may well have evolved long before the start of the Cambrian. It is also difficult to understand why the evolution of eyesight would have caused an explosion, since other senses such as smell and pressure detection can detect things further away than they can be seen under the sea, but the appearance of these other senses apparently did not cause an evolutionary explosion.

Arms races between predators and prey

The ability to avoid or recover from predation often makes the difference between life and death, and is therefore one of the strongest components of natural selection. The pressure to adapt is stronger on the prey than on the predator: if the predator fails to win a contest, it loses its lunch; if the prey is the loser, it loses its life.

But there is evidence that predation was rife long before the start of the Cambrian, for example in the increasingly spiny forms of acritarchs, the holes drilled in *Cloudina* shells, and traces of burrowing to avoid predators. Hence it is unlikely that the *appearance* of predation was the trigger for the Cambrian "explosion", although it may well have exhibited a strong influence on the body forms that the "explosion" produced. However, the intensity of predation does appear to have increased dramatically during the Cambrian as new predatory "tactics" (such as shell-crushing) emerged.

Increase in size and diversity of planktonic animals

Geochemical evidence strongly indicates that the total mass of plankton has been similar to modern levels since early in the Proterozoic. Before the start of the Cambrian, their corpses and droppings were too small to fall quickly towards the seabed, since their drag was about the same as their weight. This meant they were destroyed by scavengers or by chemical processes before they reached the sea floor.

Mesozooplankton are plankton of a larger size, and early Cambrian specimens filtered microscopic plankton from the seawater. These larger organisms would have produced droppings and corpses that were large enough to fall fairly quickly. This provided a new supply of energy and nutrients to the mid-levels and bottoms of the seas, which opened up a huge range of new possible ways of life. If any of these remains sank uneaten to the sea floor they could be buried; this would have taken some carbon out of circulation, resulting in an increase in the concentration of breathable oxygen in the seas (carbon readily combines with oxygen).

The initial herbivorous mesozooplankton were probably larvae of benthic (seafloor) animals. A larval stage was probably an evolutionary innovation driven by the increasing level of predation at the seafloor during the Ediacaran period.

Metazoans have an amazing ability to increase diversity through co-evolution. This means that a trait of one organism can cause another to evolve in response; a number of responses are possible, and a different species can potentially emerge for each. As a simple example, the evolution of predation may have caused one organism to develop defence while another developed motion to flee. This would cause the predator lineage to split into two species: one that was good at chasing prey, and another that was good at breaking through defences. Actual co-evolution is somewhat more subtle, but in this fashion, great diversity can arise: three quarters of living species are animals, and most of the rest have formed by co-evolution with animals.

Mis-dating of species

The results of an article published in Nature in 2010, have shown that eukaryotic multicellularity, which had been thought to evolve with the beginning of Cambrian Period, might date back to 2.1 billion years ago, which is approximately 1.55 billion years earlier than the date indicated by currently dominating scientific evidence.

Discredited hypotheses

As our understanding of the events of the Cambrian becomes clearer, data has accumulated to make some hypotheses look improbable. Causes that have been proposed but are now discounted include the evolution of herbivory, vast changes in the speed of tectonic plate movement or of the cyclic changes in the Earth's orbital motion, or the operation of different evolutionary mechanisms from those that are seen in the rest of the Phanerozoic eon.

No explanation required

The explosion may not have been a significant evolutionary event. It may represent a threshold being crossed: for example a threshold in genetic complexity that allowed a vast range of morphological forms to be employed.

Uniqueness of the explosion

The "Cambrian explosion" can be viewed as two waves of metazoan expansion into empty niches: first, a co-evolutionary rise in diversity as animals explored niches on the

Ediacaran sea floor, followed by a second expansion in the early Cambrian as they became established in the water column. The rate of diversification seen in the Cambrian phase of the explosion is unparalleled among marine animals: it affected all metazoan clades of which Cambrian fossils have been found. Later radiations, such as those of fish in the Silurian and Devonian periods, involved fewer taxa, mainly with very similar body plans. Although the recovery from the Permian-Triassic extinction started with about as few animal species as the Cambrian explosion, the recovery produced far fewer significantly new types of animals.

Whatever triggered the early Cambrian diversification opened up an exceptionally wide range of previously-unavailable ecological niches. When these were all occupied, there was little room for such wide-ranging diversifications to occur again, because there was strong competition in all niches and incumbents usually had the advantage. If there had continued to be a wide range of empty niches, clades would be able to continue diversifying and become disparate enough for us to recognise them as different phyla; when niches are filled, lineages will continue to resemble one another long after they diverge, as there is limited opportunity for them to change their life-styles and forms.

There were two similar explosions in the evolution of land plants: after a cryptic history beginning about 450 million years ago, land plants underwent a uniquely rapid adaptive radiation during the Devonian period, about 400 million years ago. Furthermore, Angiosperms (flowering plants) originated and rapidly diversified during the Cretaceous period.

Chapter 5

Lepidoptera Migration

Monarch butterfly is the most well known migrant amongst the Lepidoptera.

Tirumala septentrionis migrate in millions between Eastern Ghats and Western Ghats in India.

Catopsilia pomona migrate regularly in South India and Himalayas.

Lepidoptera migration is a biological phenomenon whereby populations of butterflies or moths migrate over long distances to areas where they cannot settle for long periods of time. The term *migratory butterfly* or moth does not indicate a taxonomic clade within the Lepidoptera, but is a term that is used for all such species from the various included families that migrate.

Lepidoptera species migrate on all continents except Antarctica; they migrate from or within subtropical and tropical areas. The terms *migratory butterfly* and *migratory moth* are location-bound: they are only known as such in the areas where the species cannot establish themselves permanently.

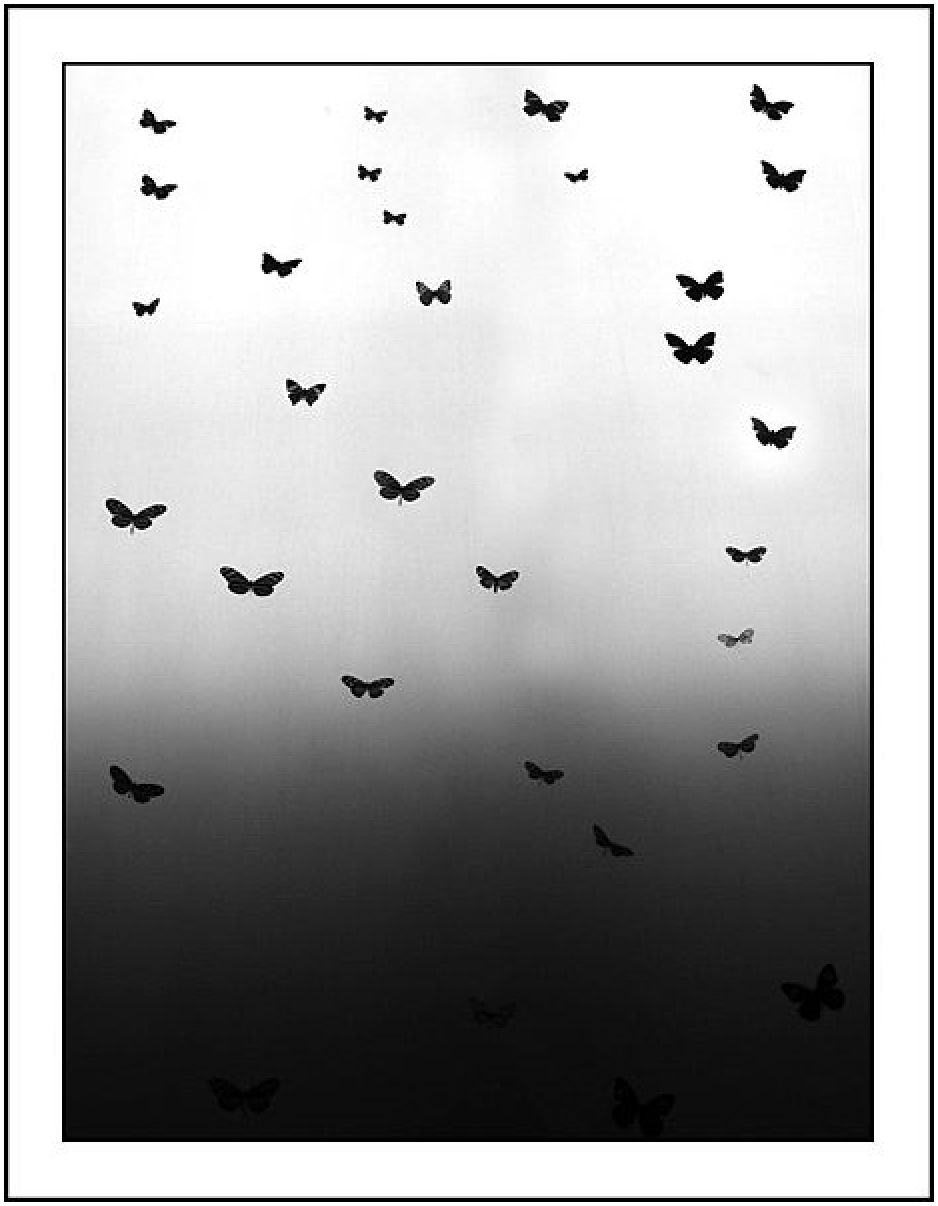
By migrating, Lepidoptera species can avoid unfavorable circumstances, including weather, food shortage, or over-population. Like birds, there are Lepidoptera species of which all individuals migrate, but there are also species of which only a subgroup of the individuals migrate.

The most famous Lepidopteran migration is that of the Monarch butterfly which migrates from southern Canada to wintering sites in central Mexico. In late winter/early spring, the adult monarchs leave the Transvolcanic mountain range in Mexico for a more northern climate. Mating occurs and the females begin seeking out milkweed to lay their eggs, usually first in northern Mexico and southern Texas. The caterpillars hatch and develop into adults that move north, where more offspring can go as far as Central Canada until next migratory cycle.

The other interesting migration is that of Danaids in South India which happens between Eastern Ghats and Western Ghats. Three species will be involved in this, namely *Tirumala septentrionis*, *Euploea core*, and *Euploea sylvester*. Sometimes they are joined by Lemon Pansy (*Junonia lemonias*), Common Emigrant (*Catopsilia pomona*), Tawny coster (*Acraea terpsicore*) and Blue Tiger (*Tirumala limniace*).

Butterfly migration happens regularly in the foot hills of Himalayas. Due to lack of continuous study in this area, we don't have substantial information.

Definition



Lepidoptera migration

Migration in Lepidoptera refers to the act of moving from one place to another in a manner that is seasonally determined and predictable. There is no unambiguous definition of *migratory butterfly* or *migratory moth*, and this also applies to proposals to divide them into classes. Migration takes place so that organisms can avoid unfavourable environmental conditions that limit breeding.

Migration means different things to behavioral scientists and ecologists. The former emphasize the act of moving whereas the latter discriminate between whether the movement has been ecologically significant or not. Migration may be viewed as "a behavioural process with ecological consequences".

Migration in Lepidoptera takes place in two of the three modes of migration identified by Johnson (Johnson, 1969). In the first mode (also Johnson's first), the Lepidoptera move in one direction in their short life-span and do not return. An example is the pierid butterfly, *Ascia monuste*, which breeds in Florida but sometimes migrates along the coast up to 160 kilometers to breed in more suitable areas.

In the second mode (Johnson's third), migration takes place to a place of hibernation or aestivation where they undergo diapause and the same generation survives to return. The classic example is that of the nymphalid Monarch butterfly (*Danuas plexippus*).

The term *migration* is sometimes also applied to species which move over a vast area on their own strength, but remain within their known habitat, although these are mostly referred to as *strays*. A well-known example of this is the Large White. This species moves in great numbers, but only to areas where they can remain permanently.

In practice, the terms *stray* and *migratory butterfly or moth* overlap; there are species that can survive in a certain area, but in too small numbers to maintain the species in that area permanently. On the other hand, there are species that form a population for more than two years, but are still considered strays. Only when a species has a permanent population in an area for ten years, it is referred to as native.

To add to the confusion, the term *stray* is also used for migratory species that are not recorded each year, but from time to time. Species that have been recorded very few times are referred to as vagrants.

Adventive

Species that are recorded in unexpected areas (adventive species) are not considered to be migratory species, because these did not leave their habitat on their own strength. Examples are species that are imported as egg or caterpillar alongside of their host plants or individuals that where reared by a collector but have escaped. An example of an introduced species is *Galleria mellonella*, which is found all over the world, because it is reared as food for captive birds and reptiles. At times it is difficult to decide if a species is adventive or migratory. Migratory species like *Chrysodeixis chalcites* and *Helicoverpa armigera* would be able to reach Western Europe on their own, but are also common in greenhouses.

Seasonal migration

Australian species *Agrotis infusa* spends the summer as an imago in the Australian Alps, and is sometimes found in buildings in large numbers.

Lepidoptera migration is often seasonal. With species of which all individuals migrate, the population moves between areas in the summer and winter season or the dry and wet season.

For species of which only part of the population migrates, seasonal migration is hard to determine. They can maintain themselves in part of their habitat but also reach areas where they cannot establish a permanent population. They only live there in the season that is most favorable for the species. Some of the species have the habit of returning to their permanent residence at the end of the season.

Difference with bird migration

An important difference with bird migration is that an individual butterfly or moth usually migrates in one direction, while birds migrate back and forth multiple times within their lifespan. This is due to the short lifespan as an imago. Species that migrate back and forth, usually do so in different generations. There are however, some exceptions:

- The famous migration of the Monarch butterfly in North America. This species migrates back and forth in one generation, though it completes only part of the journey in both directions in that generation. No individual completes the entire journey which is spread over a number of generations. The imago of the last summer generation is born in North America, migrates to Mexico, Florida, or California and stays there for the winter. After the winter it migrates back to the north to reproduce. In a couple of generations, the monarch migrates north to Canada.
- The migration of *Agrotis infusa* in Australia. This species migrates from south-eastern Australia to the Australian Alps in the summer to avoid the heat. After the summer it returns to reproduce.

Flight behaviour

The small Diamondback moth is a migratory species which can be found at altitudes of a 100 meters. It is capable of covering over 3,000 kilometers.

Migratory Lepidoptera are, in most cases, excellent flyers. Species like the *Vanessa atalanta* are capable of managing a fierce headwind. In case of headwind, they usually fly low and are more goal-oriented. During migration, some species can be found on high altitudes, ranging to up to two kilometers This is especially noteworthy for day-flying species like *Vanessa atalanta*, since the temperatures on these altitudes are low and day-flying species depend on the outside temperature to stay warm. It is thought that *Vanessa atalanta* produces enough body warmth during flight since it has also been recorded migrating at night.

In the case of transcontinental migration where distances are large, the flying speed of the butterfly (of the order of 3 metres per second or less) is inadequate for timely completion of journey. The migration is carried out by relying on heavy winds; a persistent wind speed of 10 metres per second being able to provide a displacement of 300 to 400 kilometers in a single day. For example, the Painted Lady (*Vanessa cardui*) migrates from Africa to Spain aided by head winds.

That the migratory species are good flyers, is not the same as saying they are robust flyers. The small Diamondback moth is also a migratory species that migrates 3,000 kilometers and can be found up to altitudes of 100 meters or more.

Navigation

To be able to migrate over long distances, species must be able to navigate. There are several ways they do this.

Landscape: Lepidoptera use coastal lines, mountains, but also man-made roads to orient themselves. Above sea it has been observed that the flight direction is much more accurate if the landscape on the coast is still visible.

Celestial navigation: Butterflies are known to be capable of navigation with the help of the sun. They can also navigate by using polarized light. The polarization of the sun's light changes with the angle of the rays, hence they can also navigate with cloudy weather. There are indications that they can even make corrections depending on the time on a day. Diamondback moths are known to fly in a straight trajectory which is not dependent on the angle of the sun's rays. Tests have been performed to interfere with the biological clock of certain species by keeping them in the dark and then observing if they would choose for other flight paths. The conclusion was that some species did, and others did not. Research on monarchs demonstrates that with removal of antennae, the location of the circadian clock, individuals do not localize in any one direction during flight as they do with antennae intact. Night flyers cannot use sun light for navigation. Most of these species rely on the moon and stars instead.

Earth’s magnetic field: A number of moths use the Earth's magnetic field to navigate, as a study of the stray Heart and Dart suggests. Another study, this time of the migratory behaviour of the Silver Y, showed that this species, even at high altitudes, can correct its course with changing winds, and prefers flying with favourable winds, which suggests a great sense of direction. *Aphrissa statira* in Panama loses its navigational capacity when exposed to a magnetic field, suggesting it uses the Earth’s magnetic field.

Areas where migratory butterflies and moths can be found

Migratory Lepidoptera can be found on all continents, migrating within or from the Tropics and Subtropics. North, they can be found up to Spitsbergen, above the Arctic Circle. Some migratory Lepidoptera have spread over most of the world. Some of these are pest insects, such as the diamondback moth, *Helicoverpa armigera*, and *Trichoplusia ni*.

Examples of migratory Lepidoptera

Some examples of Migratory Lepidoptera are:

- In the Indian subcontinent, migrating Lepidoptera are common just before the monsoon season. Over 250 species are known to migrate during this period.
- On Madagascar, *Chrysiridia rhipheus* migrates between populations of four plant species from the *Omphalea* genus, the host plant of this species. The three western *Omphalea* species live in dry coniferous forests, the eastern species is found in the rainforest and is the only species that is green year round.
- *Rhodometra sacraria* is normally found in Africa, large parts of Asia and Southern Europe. At times, it migrates north and can reach central and northern Europe.
- *Vanessa cardui* is found all over the world, except South America, on altitudes of up to 3,000 meters above sea level. Its home, however, are subtropical steppe areas.
- Several South and Central American species of *Uraniidae* display a great peak in migratory behavior in certain years. In the years with a great number of migrating individuals, there are “population explosions”. Individuals migrate to the south and east. There is no real re-migration to speak off. These moths feed on *Omphalea* species. These can be found in scattered populations all across South and Central America, but only part of the areas where they are found are a permanent habitat for these moths. Research suggests that the cause of the migration peak is an increase in toxicity when much of a single plant is eaten and decreasing toxicity when only small amounts of a certain plant are eaten.
- The migration of *Euploea core*, *Euploea sylvester* and *Tirumala septentrionis* between Western Ghats and Eastern Ghats in India covering up to 350 to 400 km. This migration happens twice a year. The most probable reasons seem to be the Monsoons, lack of diapause and host plants, nectar and alkaloidal resource availability.

Tirumala septentrionis

Vanessa cardui

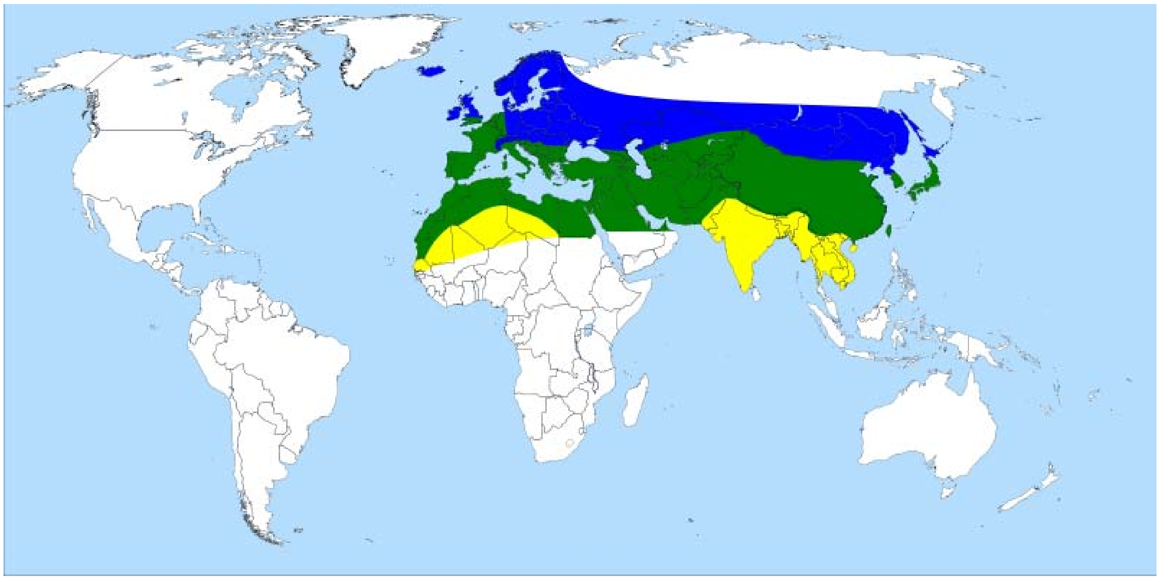
Appias albina

Catopsilia pomona

Euploea core

Euploea sylvester

Example species: *Macroglossum stellatarum*



The green areas are the areas where *Macroglossum stellatarum* can survive year round. In the blue areas it can be found in summer. In the yellow area it is found in winter

Macroglossum stellatarum

Macroglossum stellatarum is a moth that is recorded in the subtropical part of the Palearctic ecozone year round. In summer, the species migrate north up to Scandinavia and Iceland. In winter it migrates further south, deeper into Africa and to the Indian subcontinent.

In the Netherlands and Belgium, there are 100 to 200 records per year in an average year. In warm summers however, like the years 2005 and 2006, several thousand are recorded. In mild winters, small numbers can survive this far north, but these numbers are insufficient to call it a real population.

Causes

Usually, butterflies and moths migrate to escape from potentially harmful circumstances. Examples of this are a shortage of proper food plants, an unfavorable climate, like cold or extreme rain or overpopulation.

Migration and evolution

Death's Head Moth (*Acherontia atropos*) females are significantly less fertile or not fertile at all in the Netherlands.

A phenomenon like migration is an evolutionary development. By migrating, the species has survived the process of natural selection. There are a number of advantages and disadvantages to migration.

Positive effects for the species

There are positive effects for all Lepidoptera species when migrating over even small distances. It ensures sufficient genetic mixing, which in turn ensures there is no inbreeding. When the displacement takes place over longer distances, this can also lead to the discovery of new suitable habitats, making the species less vulnerable.

Furthermore, displacement over long distances makes it possible to escape from unfavorable circumstances like drought or cold. This makes migration a suitable alternative for diapause. After the negative circumstances have passed, members of the species return to their place of origin.

There are also migratory species that move over long distances, without any recorded re-migration, like several "Sphingidae" species. Some individuals possibly return to their original habitat, but these are only small numbers. The advantages of migration are not all that clear for these species. Researchers think the causes are the food plants of these species, which are mainly herbaceous plants. These plants are more vulnerable for external influences, like drought, in comparison with shrubs or trees. To tackle this problem, a possible solution is migration, which has the additional benefit for the species also escaping from these negative external influences. The fact that many of the migrating individuals find themselves in areas where they cannot survive permanently is apparently of less significance. Furthermore, even a small number of re-migrating individuals can ensure the re-emerging of a population in their original habitat.

Negative effects for the species

Migration has some negative effects on a species. The migration itself costs a lot of energy and is not without risk. During the trip, many individuals die and not all survivors find a suitable habitat. On the outskirts of suitable habitat, fertility can decrease. *Acherontia atropos* females that are born in temperate areas are significantly less fertile, or not fertile at all.

Reproduction in favorable areas must thus be very high, to shift the balance in favor of migration.

Recording

Abbildung von einer unzehligen Menge frembden großen geflügelten Ungeziefſers, welches bey Turnau und angräntzenden Gebürg im Eingang Monaths *Septembris* 1719. gefallen. Denen Garten u: Feld Früchten großen Schaden zugefüget, doch aber ſich bald wieder verlohren haben, wor bey meiſt zu *conſideriren*, daß ſie über ihren Maul einen zuſam̄en gerollten dünnen Rüſſel bey $1\frac{1}{2}$ Finger lang gehabt, den ſie gleich einer langen Nadel auß ſtrecken kön̄en, was dieſe zu deuten, iſt Gott bekannt.

Record of a massive appearance of (presumably) *Agrius convolvuli* in 1719 near Turnau.

Butterflies (and to a lesser extent moths) migrating in large numbers are a noteworthy sight, which is easily to observe and track. There are several historic records about migrating butterflies. There are records dating back to 1100 about migrating butterflies (probably a *Pieris* species) from Bavaria to the Duchy of Saxony and from 1248 about the migration of yellow butterflies in Japan.

When flying at high altitudes, spotting migrating butterflies or moths can be hard. Low flying species are easily spotted or caught using a light trap. When individuals fly too high for these methods, air balloons equipped with nets are used at times. Alternatively, radar is used to monitor migration.

Another registration technique, is marking the wings with tiny stickers, a technique comparable with Bird ringing. This technique has not proven to be very successful though. Advances in technology might make it possible to equip individuals with micro transmitters in the future.

Registration

Both The Netherlands and Belgium have a register of recorded migratory species. In the Netherlands, registration (*Trekvlinderregistratie Nederland*) started in 1940 and is the world's oldest project in this field., in Belgium, the registration is called *Belgisch Trekvlinder Onderzoek*.

Migration and climate change

It seems that global warming has caused an increase of migratory butterflies and moths that reach north-western countries like the Netherlands, Belgium and the United Kingdom. Research in the United Kingdom confirms that an increasing number of migrants reach the country. Because one would expect that migratory species can adapt to new circumstances quite well, the researches warn for new species that can have a negative impact on native species and possible damage to both health (species like the Oak Processionary) and agriculture.

Chapter 6

Evolution of Sexual Reproduction

The **evolution of sexual reproduction** is currently described by several competing scientific hypotheses. Many groups of organisms, notably the majority of animals and plants, reproduce sexually. The evolution of sex contains two related, yet distinct, themes: its *origin* and its *maintenance*. However, since the hypotheses for the origins of sex are difficult to test experimentally, most current work has been focused on the maintenance of sexual reproduction.

It seems that a sexual cycle is maintained because it improves the quality of progeny (fitness), despite reducing the overall number of offspring (the twofold cost of sex). In order for sex to be evolutionarily advantageous, it must be associated with a significant increase in the fitness of offspring. One of the most widely accepted explanations for the advantage of sex lies in the *creation of genetic variation*. Another explanation is based on two molecular advantages. First is the advantage of *recombinational DNA repair* (promoted during meiosis because homologous chromosomes pair at that time), and second is the advantage of *complementation* (also known as hybrid vigor, heterosis or masking of mutations).

For the advantage due to creation of genetic variation, there are three possible reasons this might happen. First, sexual reproduction can bring together mutations that are beneficial into the same individual (*sex aids in the spread of advantageous traits*). Second, sex acts to bring together currently deleterious mutations to create severely unfit individuals that are then eliminated from the population (*sex aids in the removal of deleterious genes*). Last, sex creates new gene *combinations* that may be more fit than previously existing ones, or may simply lead to reduced competition among relatives.

For the advantage due to *DNA repair*, there is an immediate large benefit to removal of DNA damage by recombinational DNA repair during meiosis, since this removal allows greater survival of progeny with undamaged DNA. The advantage of *complementation* to each sexual partner is avoidance of the bad effects of their deleterious recessive genes in progeny by the masking effect of normal dominant genes contributed by the other partner.

The classes of hypotheses based on the *creation of variation* are further broken down below. It is important to realise that any number of these hypotheses may be true in any given species (they are not mutually exclusive), and that different hypotheses may apply in different species. However, a research framework based on *creation of variation* has yet to be found that allows one to determine whether the reason for sex is universal for all sexual species, and, if not, which mechanism is acting in each species.

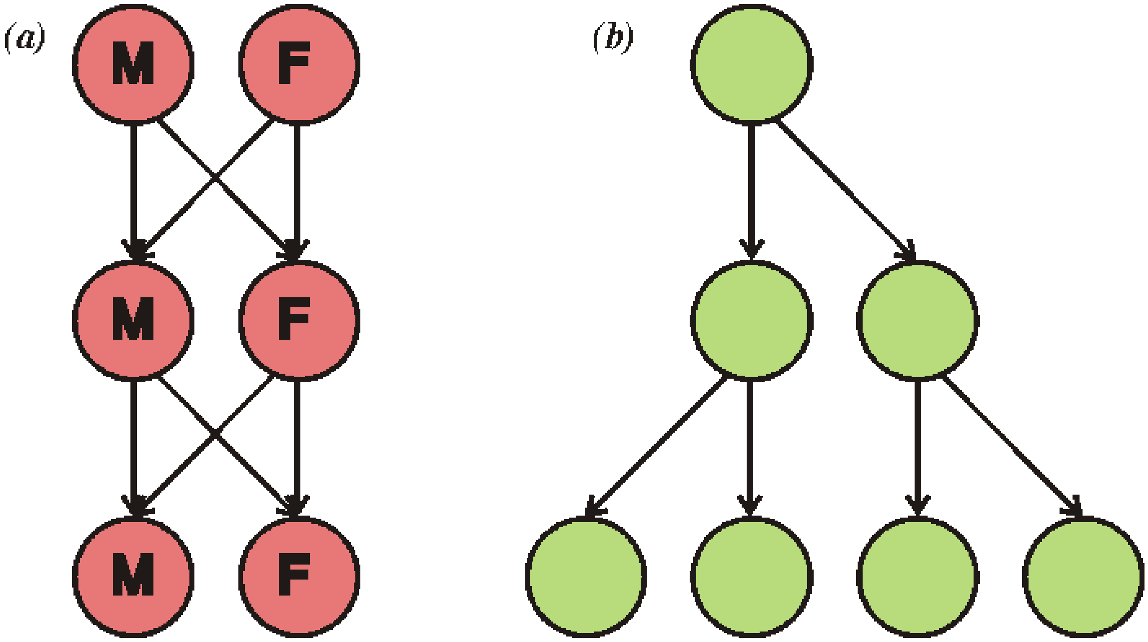
On the other hand, the maintenance of sex based on DNA repair and complementation applies widely to all sexual species. This explanation for the maintenance of sex is explored further in Section 6.2.

Historical perspective

Modern philosophical-scientific thinking on the problem can be traced back to Erasmus Darwin in the 18th century; it also features in Aristotle's writings. The thread was later picked up by August Weismann in 1889, who understood that the purpose of sex was to generate genetic variation, as is detailed in the majority of the explanations below. On the other hand, Charles Darwin, also in 1889, concluded that the effects of hybrid vigor (complementation) "is amply sufficient to account for the ... genesis of the two sexes." This is consistent with the repair and complementation hypothesis, given below under "Other explanations."

Several explanations have been suggested by biologists including W. D. Hamilton, Alexey Kondrashov, George C. Williams, Harris Bernstein, Carol Bernstein, Michael M. Cox, Frederic A. Hopf and Richard E. Michod to explain how sexual reproduction is maintained in a vast array of different living organisms.

Two-fold cost of sex



This diagram illustrates the *twofold cost of sex*. If each individual were to contribute to the same number of offspring (two), *(a)* the sexual population remains the same size each generation, where the *(b)* asexual population doubles in size each generation.

In all multicellular sexual species, the population consists of two sexes, only one of which is capable of bearing young (with the exception of simultaneous hermaphrodites). In an asexual species, each member of the population is capable of bearing young. This implies that an asexual population has an intrinsic capacity to grow more rapidly each generation. The cost was first described in mathematical terms by John Maynard Smith. He imagined an asexual mutant arising in a sexual population, half of which comprises males that cannot themselves produce offspring. With female-only offspring, the asexual lineage doubles its representation in the population each generation, all else being equal. Often all else is not equal, however, in which case the realised fitness cost to sex may be much less than this intrinsic twofold cost of producing males. For example, an asexual mutant arising in a sexual population occupies a niche frozen to that of its parental genotype because the asexual descendants are genetically self-identical. Analysis of competitive Lotka-Volterra equations suggests that the asexual lineage may never realise its full twofold advantage in population growth capacity, if the broader niche of the sexual population confers even a small competitive edge.

An additional cost is that males and females must search for each other in order to mate, and sexual selection often favours traits that reduce the survival of individuals.

Evidence that the cost is not insurmountable comes from George C. Williams, who noted the existence of species which are capable of both asexual and sexual reproduction. These species time their sexual reproduction with periods of environmental uncertainty, and reproduce asexually when conditions are more favourable. The important point is that these species are observed to reproduce sexually when they could choose not to, implying that there is a selective advantage to sexual reproduction.

Another disadvantage of sexual reproduction is that a sexually reproducing organism will only be able to pass on 50% of its genes to each offspring. This is a consequence of the fact that gametes from sexually reproducing species are haploid.

Promotion of genetic variation

August Weismann proposed in 1889 an explanation for the evolution of sex, where the advantage of sex is the creation of variation among siblings. It was then subsequently explained in genetics terms by Fisher and Muller and has been recently summarised by Burt in 2000.

George C. Williams gave an example based around the elm tree. In the forest of this example, empty patches between trees can support one individual each. When a patch becomes available because of the death of a tree, other trees' seeds will compete to fill the patch. Since the chance of a seed's success in occupying the patch depends upon its genotype, and a parent cannot anticipate which genotype is most successful, each parent will send many seeds, creating competition between siblings. Natural selection therefore favours parents which can produce a variety of offspring.

A similar hypothesis is named the *tangled bank hypothesis* after a passage in Charles Darwin's *The Origin of Species*:

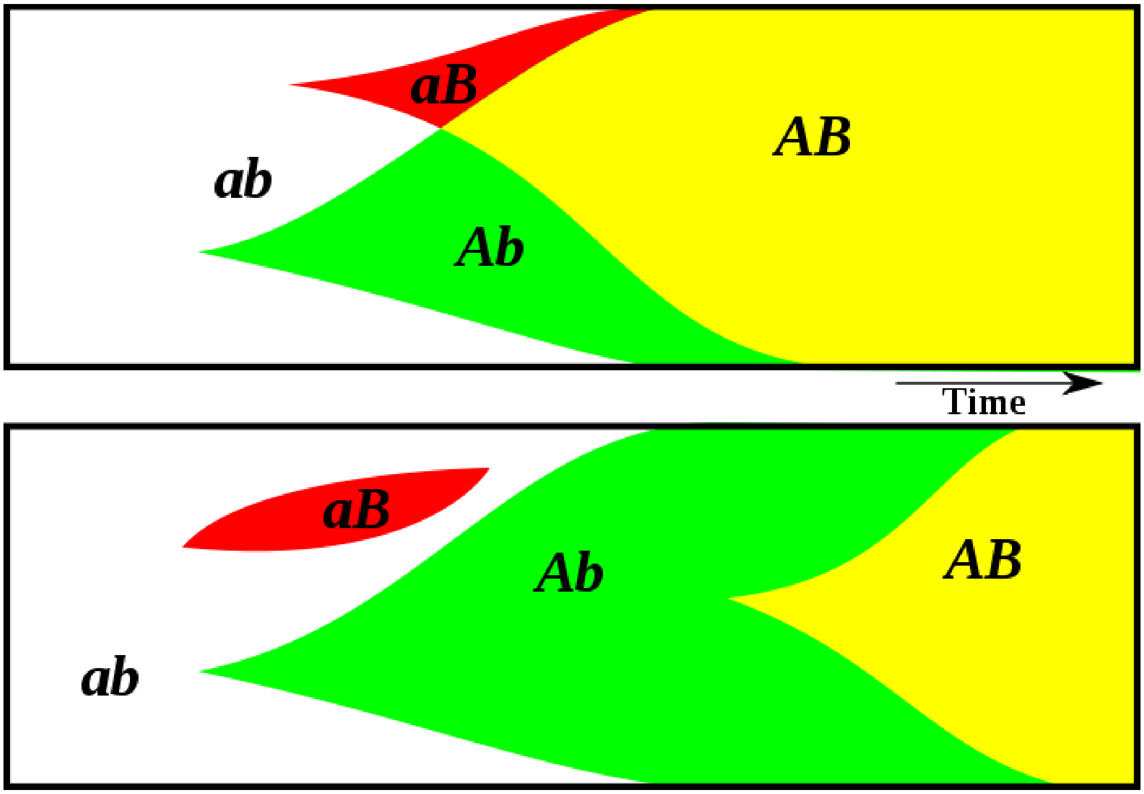
"It is interesting to contemplate an entangled bank, clothed with many plants of many kinds, with birds singing on the bushes, with various insects flitting about, and with worms crawling through the damp earth, and to reflect that these elaborately constructed forms, so different from each other, and dependent on each other in so complex a manner, have all been produced by laws acting around us."

The hypothesis, proposed by Michael Ghiselin in his 1974 book, *The Economy of Nature and the Evolution of Sex*, suggests that a diverse set of siblings may be able to extract more food from its environment than a clone, because each sibling uses a slightly different niche. One of the main proponents of this hypothesis is Graham Bell of McGill University. The hypothesis has been criticised for failing to explain how asexual species developed sexes. In his book, *Evolution and Human Behavior* (MIT Press, 2000), John Cartwright comments:

"Although once popular, the tangled bank hypothesis now seems to face many problems, and former adherents are falling away. The theory would predict a greater interest in sex among animals that produce lots of small offspring that compete with each other. In fact, sex is invariably associated with organisms that produce a few large offspring, whereas organisms producing small offspring frequently engage in parthenogenesis [asexual reproduction]. In addition, the evidence from fossils suggests that species go for vast periods of [geologic] time without changing much."

Spread of advantageous traits

Novel genotypes



This diagram illustrates how sex might create novel genotypes more rapidly. Two advantageous alleles *A* and *B* occur at random. The two alleles are recombined rapidly in a sexual population (top), but in an asexual population (bottom) the two alleles must independently arise because of clonal interference.

Sex could be a method by which novel genotypes are created. Since sex combines genes from two individuals, sexually reproducing populations can more easily combine advantageous genes than can asexual populations. If, in a sexual population, two different advantageous alleles arise at different loci on a chromosome in different members of the population, a chromosome containing the two advantageous alleles can be produced within a few generations by recombination. However, should the same two alleles arise in different members of an asexual population, the only way that one chromosome can develop the other allele is to independently gain the same mutation, which would take much longer.

Ronald Fisher also suggested that sex might facilitate the spread of advantageous genes by allowing them to escape their genetic surroundings, if they should arise on a chromosome with deleterious genes.

Supporters of these theories respond to the balance argument that the individuals produced by sexual and asexual reproduction may differ in other respects too – which may influence the persistence of sexuality. For example, in water fleas of the genus *Cladocera*, sexual offspring form eggs which are better able to survive the winter.

Increased resistance to parasites

One of the most widely accepted theories to explain the persistence of sex is that it is maintained to assist sexual individuals in resisting parasites, also known as the Red Queen's Hypothesis.

When an environment changes, previously neutral or deleterious alleles can become favourable. If the environment changed sufficiently rapidly (i.e. between generations), these changes in the environment can make sex advantageous for the individual. Such rapid changes in environment are caused by the co-evolution between hosts and parasites.

Imagine, for example that there is one gene in parasites with two alleles *p* and *P* conferring two types of parasitic ability, and one gene in hosts with two alleles *h* and *H*, conferring two types of parasite resistance, such that parasites with allele *p* can attach themselves to hosts with the allele *h*, and *P* to *H*. Such a situation will lead to cyclic changes in allele frequency - as *p* increases in frequency, *h* will be disfavoured.

In reality, there will be several genes involved in the relationship between hosts and parasites. In an asexual population of hosts, offspring will only have the different parasitic resistance if a mutation arises. In a sexual population of hosts, however, offspring will have a new combination of parasitic resistance alleles.

In other words, like Lewis Carroll's Red Queen, sexual hosts are continually adapting in order to stay ahead of their parasites.

Evidence for this explanation for the evolution of sex is provided by comparison of the rate of molecular evolution of genes for kinases and immunoglobulins in the immune system with genes coding other proteins. The genes coding for immune system proteins evolve considerably faster.

Critics of the Red Queen hypothesis question whether the constantly-changing environment of hosts and parasites is sufficiently common to explain the evolution of sex.

Deleterious mutation clearance

Mutations can have many different effects upon an organism. It is generally believed that the majority of non-neutral mutations are deleterious, which means that they will cause a decrease in the organism's overall fitness. If a mutation has a deleterious effect, it will then usually be removed from the population by the process of natural selection. Sexual reproduction is believed to be more efficient than asexual reproduction in removing those mutations from the genome.

There are two main hypotheses which explain how sex may act to remove deleterious genes from the genome.

Maintenance of mutation-free individuals

In a finite asexual population under the pressure of deleterious mutations, the random loss of individuals without such mutations is inevitable. This is known as Muller's ratchet. In a sexual population, however, mutation-free genotypes can be restored by recombination of genotypes containing deleterious mutations.

This comparison will only work for a small population - in a large population, random loss of the most fit genotype becomes unlikely even in an asexual population.

Removal of deleterious genes

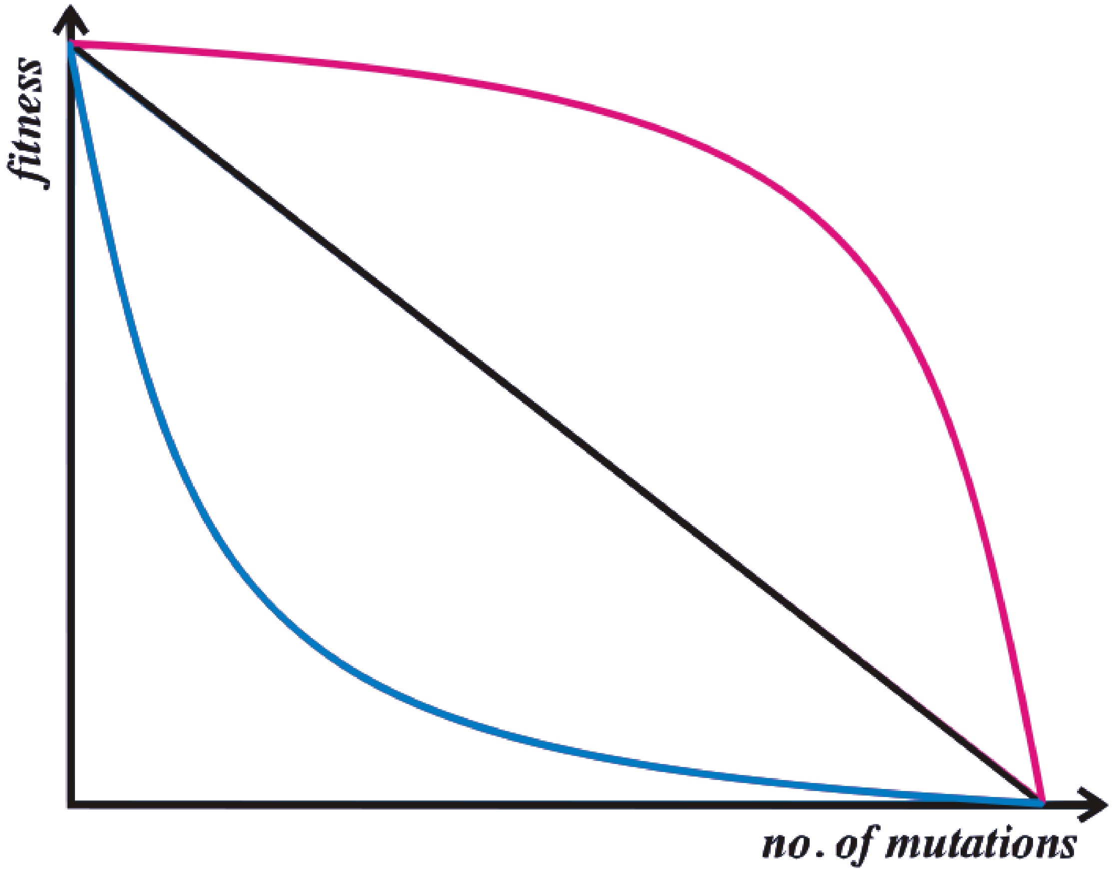


Diagram illustrating different relationships between numbers of mutations and fitness. Kondrashov's model requires *synergistic epistasis*, which is represented by the blue line - each mutation has a disproproportionately large effect on the organism's fitness.

This hypothesis was proposed by Alexey Kondrashov, and is sometimes known as the *deterministic mutation hypothesis*. It assumes that the majority of deleterious mutations are only slightly deleterious, and affect the individual such that the introduction of each additional mutation has an increasingly large effect on the fitness of the organism. This relationship between number of mutations and fitness is known as *synergistic epistasis*.

By way of analogy, think of a car with several minor faults. Each is not sufficient alone to prevent the car from running, but in combination, the faults combine to prevent the car from functioning.

Similarly, an organism may be able to cope with a few defects, but the presence of many mutations could overwhelm its backup mechanisms.

Kondrashov argues that the slightly deleterious nature of mutations means that the population will tend to be composed of individuals with a small number of mutations. Sex will act to recombine these genotypes, creating some individuals with fewer deleterious mutations, and some with more. Because there is a major selective disadvantage to individuals with more mutations, these individuals die out. In essence, sex compartmentalises the deleterious mutations.

There has been much criticism of Kondrashov's theory, since it relies on two key restrictive conditions. The first requires that the rate of deleterious mutation should exceed one per genome per generation in order to provide a substantial advantage for sex. While there is some empirical evidence for it (for example in Drosophila and E. coli), there is also strong evidence against it. Secondly, there should be strong interactions among loci (synergistic epistasis), a mutation-fitness relation for which there is only limited evidence. Conversely, there is also the same amount of evidence that mutations show no epistasis (purely additive model) or antagonistic interactions (each additional mutation has a disproportionally *small* effect).

Other explanations

Speed of evolution

Ilan Eshel suggested that sex prevents rapid evolution. He suggests that recombination breaks up favourable gene combinations more often than it creates them, and sex is maintained because it ensures selection is longer-term than in asexual populations - so the population is less affected by short-term changes. This explanation is not widely accepted, as its assumptions are very restrictive.

It has recently been shown in experiments with *Chlamydomonas* algae that sex can remove the speed limit on evolution.

DNA repair and complementation

As discussed in the earlier, sexual reproduction is conventionally explained as an adaptation for producing genetic variation through allelic recombination. As acknowledged above, however, serious problems with this explanation have led many biologists to conclude that the benefit of sex is a major unsolved problem in evolutionary biology.

An alternative "informational" approach to this problem has led to the view that the two fundamental aspects of sex, genetic recombination and outcrossing, are adaptive responses to the two major sources of "noise" in transmitting genetic information. Genetic noise can occur as either physical damage to the genome (e.g. chemically altered bases of DNA or breaks in the chromosome) or replication errors (mutations) This alternative view is referred to as the repair and complementation hypothesis, to distinguish it from the traditional variation hypothesis.

The repair and complementation hypothesis assumes that genetic recombination is fundamentally a DNA repair process, and that when it occurs during meiosis it is an adaptation for repairing the genomic DNA which is passed on to progeny. Recombinational repair is the only repair process known which can accurately remove double-strand damages in DNA, and such damages are both common in nature and ordinarily lethal if not repaired. Recombinational repair is prevalent from the simplest viruses to the most complex multicellular eukaryotes. It is effective against many different types of genomic damage, and in particular is highly efficient at overcoming double-strand damages. Studies of the mechanism of meiotic recombination indicate that meiosis is an adaptation for repairing DNA. These considerations form the basis for the first part of the repair and complementation hypothesis.

In some lines of descent from the earliest organisms, the diploid stage of the sexual cycle, which was at first transient, became the predominant stage, because it allowed complementation — the masking of deleterious recessive mutations (i.e. hybrid vigor or heterosis). Outcrossing, the second fundamental aspect of sex, is maintained by the advantage of masking mutations and the disadvantage of inbreeding (mating with a close relative) which allows expression of recessive mutations (commonly observed as inbreeding depression). This is in accord with Charles Darwin, who concluded that the adaptive advantage of sex is hybrid vigor; or as he put it, "the offspring of two individuals, especially if their progenitors have been subjected to very different conditions, have a great advantage in height, weight, constitutional vigor and fertility over the self fertilised offspring from either one of the same parents."

However, outcrossing may be abandoned in favor of parthogenesis or selfing (which retain the advantage of meiotic recombinational repair) under conditions in which the costs of mating are very high. For instance, costs of mating are high when individuals are rare in a geographic area, such as when there has been a forest fire and the individuals entering the burned area are the initial ones to arrive. At such times mates are hard to find, and this favors parthenogenic species.

In the view of the repair and complementation hypothesis, the removal of DNA damage by recombinational repair produces a new, less deleterious form of informational noise, allelic recombination, as a by-product. This lesser informational noise generates genetic variation, viewed by some as the major effect of sex, as discussed in the earlier here.

Origin of sexual reproduction

All sexually reproducing organisms derive from a common ancestor which was a single celled eukaryotic species. Many protists reproduce sexually, as do the multicellular plants, animals, and fungi. There are a few species which have secondarily lost this feature, such as Bdelloidea and some parthenocarpic plants.

Organisms need to replicate their genetic material in an efficient and reliable manner. The necessity to repair genetic damage is one of the leading theories explaining the origin of sexual reproduction. Diploid individuals can repair a mutated section of its DNA via homologous recombination, since there are two copies of the gene in the cell and one copy is presumed to be undamaged. A mutation in an haploid individual, on the other hand, is more likely to become resident, as the DNA repair machinery has no way of knowing what the original undamaged sequence was. The most primitive form of sex may have been one organism with damaged DNA replicating an undamaged strand from a similar organism in order to repair itself.

Another theory is that sexual reproduction originated from selfish parasitic genetic elements that exchange genetic material (that is: copies of their own genome) for their transmission and propagation. In some organisms, sexual reproduction has been shown to enhance the spread of parasitic genetic elements (e.g.: yeast, filamentous fungi). Bacterial conjugation, a form of genetic exchange that some sources describe as sex, is not a form of reproduction, but rather an example of horizontal gene transfer. However, it does support the selfish genetic element theory, as it is propagated through such a "selfish gene", the F-plasmid. Similarly, it has been proposed that sexual reproduction evolved from ancient haloarchaea through a combination of jumping genes, and swapping plasmids.

A third theory is that sex evolved as a form of cannibalism. One primitive organism ate another one, but rather than completely digesting it, some of the 'eaten' organism's DNA was incorporated into the 'eater' organism.

Sex may also be derived from prokaryotic processes. A comprehensive 'origin of sex as vaccination' theory proposes that eukaryan sex-as-syngamy (fusion sex) arose from prokaryan unilateral sex-as-infection when infected hosts began swapping nuclearised genomes containing coevolved, vertically transmitted symbionts that provided protection against horizontal superinfection by more virulent symbionts. Sex-as-meiosis (fission sex) then evolved as a host strategy to uncouple (and thereby emasculate) the acquired symbiont genomes.

Mechanistic origin of sexual reproduction

Though theories positing benefits that lead to the origin of sex are often problematic, several credible theories on the evolution of the mechanisms of sexual reproduction have been proposed.

Viral eukaryogenesis

The viral eukaryogenesis (VE) theory proposes that eukaryotic cells arose from a combination of a lysogenic virus, an archaeon and a bacterium. This model suggests that the nucleus originated when the lysogenic virus incorporated genetic material from the archaeon and the bacterium and took over the role of information storage for the amalgam. The archaeal host transferred much of its functional genome to the virus during the evolution of cytoplasm but retained the function of gene translation and general metabolism. The bacterium transferred most of its functional genome to the virus as it transitioned into a mitochondrion.

For these transformations to lead to the eukaryotic cell cycle, the VE hypothesis specifies a pox-like virus as the lysogenic virus. A pox-like virus is a likely ancestor because of its fundamental similarities with eukaryotic nuclei. These include a double stranded DNA genome, a linear chromosome with short telomeric repeats, a complex membrane bound capsid, the ability to produce capped mRNA, and the ability to export the capped mRNA across the viral membrane into the cytoplasm. The presence of a lysogenic pox-like virus ancestor explains the development of meiotic division, an essential component of sexual reproduction.

Meiotic division in the VE hypothesis arose because of the evolutionary pressures placed on the lysogenic virus as a result of its inability to enter into the lytic cycle. This selective pressure resulted in the development of processes allowing the viruses to spread horizontally throughout the population. The outcome of this selection was cell-to-cell fusion. (This is distinct from the conjugation methods used by bacterial plasmids under evolutionary pressure, with important consequences.) The possibility of this kind of fusion is supported by the presence of fusion proteins in the envelopes of the pox viruses that allow them to fuse with host membranes. These proteins could have been transferred to the cell membrane during viral reproduction, enabling cell-to-cell fusion between the virus host and an uninfected cell. The theory proposes meiosis originated from the fusion between two cells infected with related but different viruses which recognised each other as uninfected. After the fusion of the two cells, incompatibilities between the two viruses result in a meiotic-like cell division.

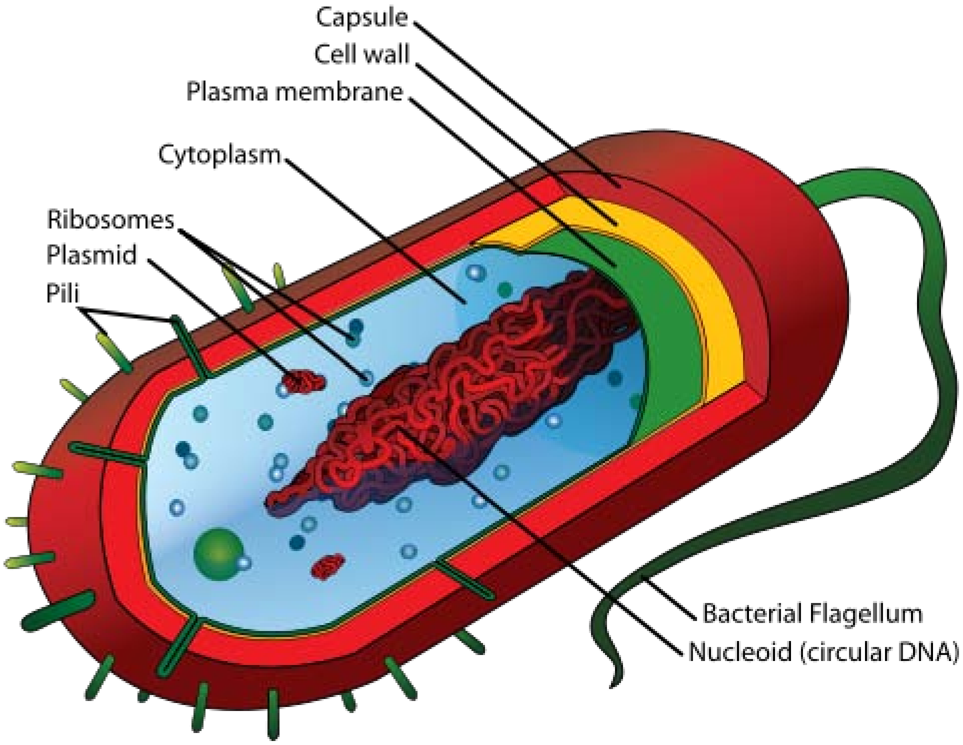
The two viruses established in the cell would initiate replication in response to signals from the host cell. A mitosis-like cell cycle would proceed until the viral membranes dissolved, at which point linear chromosomes would be bound together with centromeres. The homologous nature of the two viral centromeres would incite the grouping of both sets into tetrads. It is speculated that this grouping may be the origin of crossing over, characteristic of the first division in modern meiosis. The partitioning apparatus of the mitotic-like cell cycle the cells used to replicate independently would then pull each set of chromosomes to one side of the cell, still bound by centromeres. These centromeres would prevent their replication in subsequent division, resulting in four daughter cells with one copy of one of the two original pox-like viruses. The process resulting from combination of two similar pox viruses within the same host closely mimics meiosis.

Neomuran revolution

An alternative theory, proposed by Thomas Cavalier-Smith, was labeled the Neomuran revolution. The designation "Neomuran revolution" refers to the appearances of the common ancestors of eukaryotes and archaea. Cavalier-Smith proposes that the first neomurans emerged 850 million years ago. Other molecular biologists assume that this group appeared much earlier, but Cavalier-Smith dismisses these claims because they are based on the "theoretically and empirically" unsound model of molecular clocks. Cavalier-Smith's theory of the Neomuran revolution has implications for the evolutionary history of the cellular machinery for recombination and sex. It suggests that this machinery evolved in two distinct bouts separated by a long period of stasis; first the appearance of recombination machinery in a bacterial ancestor which was maintained for 3 Gy, until the neomuran revolution when the mechanics were adapted to the presence of nucleosomes. The archaeal products of the revolution maintained recombination machinery that was essentially bacterial, whereas the eukaryotic products broke with this bacterial continuity. They introduced cell fusion and ploidy cycles into cell life histories. Cavalier-Smith argues that both bouts of mechanical evolution were motivated by similar selective forces: the need for accurate DNA replication without loss of viability.

Chapter 7

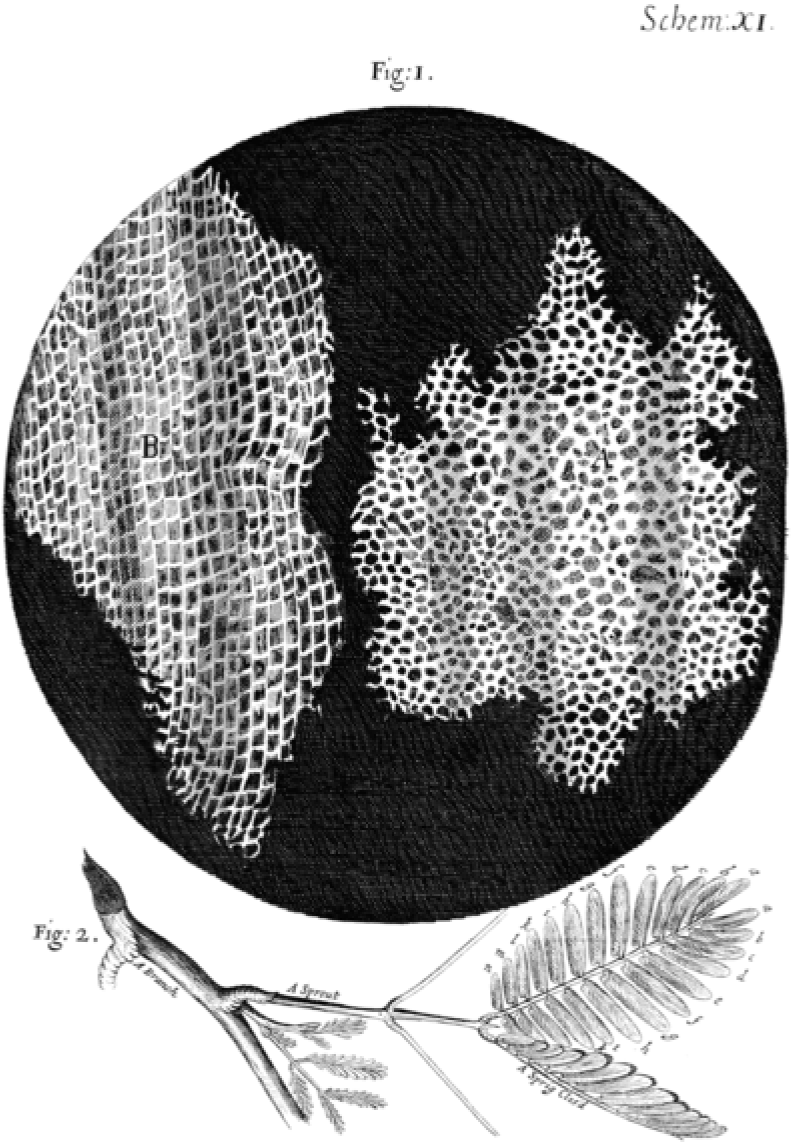
Cell Theory



A prokaryote

Cell theory refers to the idea that cells are the basic unit of structure in every living thing. Development of this theory during the mid 17th century was made possible by advances in microscopy. This theory is one of the foundations of biology. The theory says that new cells are formed from other existing cells, and that the cell is a fundamental unit of structure, function and organization in all living organisms.

History



Drawing of the structure of cork by Robert Hooke that appeared in *Micrographia*

The cell was discovered by Robert Hooke in 1665. He examined (under a coarse, compound microscope) very thin slices of cork and saw a multitude of tiny pores that he remarked looked like the walled compartments a monk would live in. Because of this association, Hooke called them cells, the name they still bear. However, Hooke did not know their real structure or function. Hooke's description of these cells (which were actually non-living cell walls) was published in *Micrographia*. His cell observations gave no indication of the nucleus and other organelles found in most living cells.

The first man to witness a live cell under a microscope was Anton van Leeuwenhoek (although the first man to make a compound microscope was Zacharias Janssen), who in 1674 described the algae *Spirogyra* and named the moving organisms animalcules, meaning "little animals". Leeuwenhoek probably also saw bacteria. Cell theory was in contrast to the vitalism theories proposed before the discovery of cells.

The idea that cells were separable into individual units was proposed by Ludolph Christian Treviranus and Johann Jacob Paul Moldenhawer. All of this finally led to Henri Dutrochet formulating one of the fundamental tenets of modern cell theory by declaring that "The cell is the fundamental element of organization"

The observations of Hooke, Leeuwenhoek, Schleiden, Schwann, Virchow, and others led to the development of the cell theory. The cell theory is a widely accepted explanation of the relationship between cells and living things. The cell theory states:

- All living things or organisms are made of cells and their products.
- New cells are created by old cells dividing into two.
- Cells are the basic building units of life.

The cell theory holds true for all living things, no matter how big or small, or how simple or complex. Since according to research, cells are common to all living things, they can provide information about all life. And because all cells come from other cells, scientists can study cells to learn about growth, reproduction, and all other functions that living things perform. By learning about cells and how they function, you can learn about all types of living things.

Credit for developing cell theory is usually given to three scientists: Theodor Schwann, Matthias Jakob Schleiden, and Rudolf Virchow. In 1839, Schwann and Schleiden suggested that cells were the basic unit of life. Their theory accepted the first two tenets of modern cell theory. However the cell theory of Schleiden differed from modern cell theory in that it proposed a method of spontaneous crystallization that he called "Free Cell Formation". In 1858, Rudolf Virchow concluded that all cells come from pre-existing cells, thus completing the classical cell theory.

Classical interpretation

1. All living organisms are made up of one or more cells.
2. Cells are the basic unit of life.
3. All cells arise from pre-existing cells.
4. The cell is the unit of structure, physiology, and organization in living things.
5. The cell retains a dual existence as a distinct entity and a building block in the construction of organisms.

Modern interpretation

The generally accepted parts of modern cell theory include:

1. The cell is the fundamental unit of structure and function in living organisms.
2. All cells arise from pre-existing cells by division.
3. Energy flow (metabolism and biochemistry) occurs within cells.
4. Cells contain hereditary information (DNA) which is passed from cell to cell during cell division.
5. All cells are basically the same in chemical composition in organisms of similar species.
6. All known living things are made up of one or more cells.
7. Some organisms are made up of only one cell and are known as unicellular organisms.
8. Others are multicellular, composed of a number of cells.
9. The activity of an organism depends on the total activity of independent cells.

Types of cells

Cells can be subdivided into the following subcategories:

1. *Prokaryotes*: Prokaryotes lack a nucleus (though they do have circular DNA) and other membrane-bound organelles (though they do contain ribosomes). Bacteria and Archaea are two domains of prokaryotes.
2. *Eukaryotes*: Eukaryotes, on the other hand, have distinct nuclei bound by a nuclear membrane and membrane-bound organelles (mitochondria, chloroplasts, lysosomes, rough and smooth endoplasmic reticulum, vacuoles). In addition, they possess organized chromosomes which store genetic material.

Chapter 8

Biodiversity

Some of the biodiversity of a coral reef

Rainforests are an example of biodiversity on the planet, and typically possess a great deal of species diversity. This is the Gambia River in Senegal's Niokolo-Koba National Park.

Biodiversity is the degree of variation of life forms within a given ecosystem, biome, or an entire planet. Biodiversity is a measure of the health of ecosystems. Greater biodiversity implies greater health. Biodiversity is in part a function of climate. In terrestrial habitats, tropical regions are typically rich whereas polar regions support fewer species.

Rapid environmental changes typically cause extinctions. One estimate is that less than 1% of the species that have existed on Earth are extant.

Since life began on Earth, five major mass extinctions and several minor events have led to large and sudden drops in biodiversity. The Phanerozoic eon (the last 540 million years) marked a rapid growth in biodiversity via the Cambrian explosion—a period during which nearly every phylum of multicellular organisms first appeared. The next 400 million years included repeated, massive biodiversity losses classified as mass extinction events. In the Carboniferous, rainforest collapse led to a great loss of plant and animal life. The Permian–Triassic extinction event, 251 million years ago, was the worst; vertebrate recovery took 30 million years. The most recent, the Cretaceous–Tertiary extinction event, occurred 65 million years ago, and has often attracted more attention than others because it resulted in the extinction of the non-avian dinosaurs.

The period since the emergence of humans has displayed an ongoing biodiversity reduction and an accompanying loss of genetic diversity. Named the Holocene extinction, the reduction is caused primarily by human impacts, particularly habitat destruction. Biodiversity's impact on human health is a major international issue.

The United Nations designated 2010 as the International Year of Biodiversity.

Etymology

The term **biological diversity** was used first by wildlife scientist and conservationist Raymond F. Dasmann in the 1968 lay book *A Different Kind of Country* advocating conservation. The term was widely adopted only after more than a decade, when in the 1980s it came into common usage in science and environmental policy. Thomas Lovejoy, in the foreword to the book **Conservation Biology**, introduced the term to the scientific community. Until then the term "natural diversity" was common, introduced by The Science Division of The Nature Conservancy in an important 1975 study, "The Preservation of Natural Diversity." By the early 1980s TNC's Science program and its head, Robert E. Jenkins, Lovejoy and other leading conservation scientists at the time in America advocated the use of "biological diversity".

The term's contracted form ***biodiversity*** may have been coined by W.G. Rosen in 1985 while planning the 1986 *National Forum on Biological Diversity* organized by the National Research Council (NRC). It first appeared in a publication in 1988 when entomologist E. O. Wilson used it as the title of the proceedings of that forum.

Since this period the term has achieved widespread use among biologists, environmentalists, political leaders, and concerned citizens.

A similar term in the United States is "natural heritage." It predates the others and is more accepted by the wider audience interested in conservation. Broader than biodiversity, it includes geology and landforms (geodiversity).

Definitions

A Sampling of fungi collected during summer 2008 in Northern Saskatchewan mixed woods, near LaRonge is an example regarding the species diversity of fungi. In this photo, there are also leaf lichens and mosses.

"Biological diversity" or "biodiversity" can have many interpretations. It is most commonly used to replace the more clearly defined and long established terms, species diversity and species richness. Biologists most often define biodiversity as the "totality of genes, species, and ecosystems of a region". An advantage of this definition is that it seems to describe most circumstances and presents a unified view of the traditional three levels at which biological variety has been identified:

- species diversity

- ecosystem diversity
- genetic diversity

In 2003 Professor Anthony Campbell at Cardiff University, UK and the Darwin Centre, Pembrokeshire, defined a fourth level: Molecular Diversity.

This multilevel construct is consistent with Dasmann and Lovejoy. An explicit definition consistent with this interpretation was first given in a paper by Bruce A. Wilcox commissioned by the International Union for the Conservation of Nature and Natural Resources (IUCN) for the 1982 World National Parks Conference. Wilcox's definition was "Biological diversity is the variety of life forms...at all levels of biological systems (i.e., molecular, organismic, population, species and ecosystem)..." The 1992 United Nations Earth Summit defined "biological diversity" as "the variability among living organisms from all sources, including, 'inter alia', terrestrial, marine, and other aquatic ecosystems, and the ecological complexes of which they are part: this includes diversity within species, between species and of ecosystems". This definition is used in the United Nations Convention on Biological Diversity.

One textbook's definition is "variation of life at all levels of biological organization".

Geneticists define it as the diversity of genes and organisms. They study processes such as mutations, gene transfer, and genome dynamics that generate evolution.

Linking biodiversity levels

Measuring diversity at one level in a group of organisms may not precisely correspond to diversity at other levels. However, tetrapod (terrestrial vertebrates) taxonomic and ecological diversity shows a very close correlation.

Distribution

A conifer forest in the Swiss Alps (National Park).

Selection bias amongst researchers may contribute to biased empirical research for modern estimates of biodiversity. In 1768 Rev. Gilbert White succinctly observed of his Selborne, Hampshire "all nature is so full, that that district produces the most variety which is the most examined."

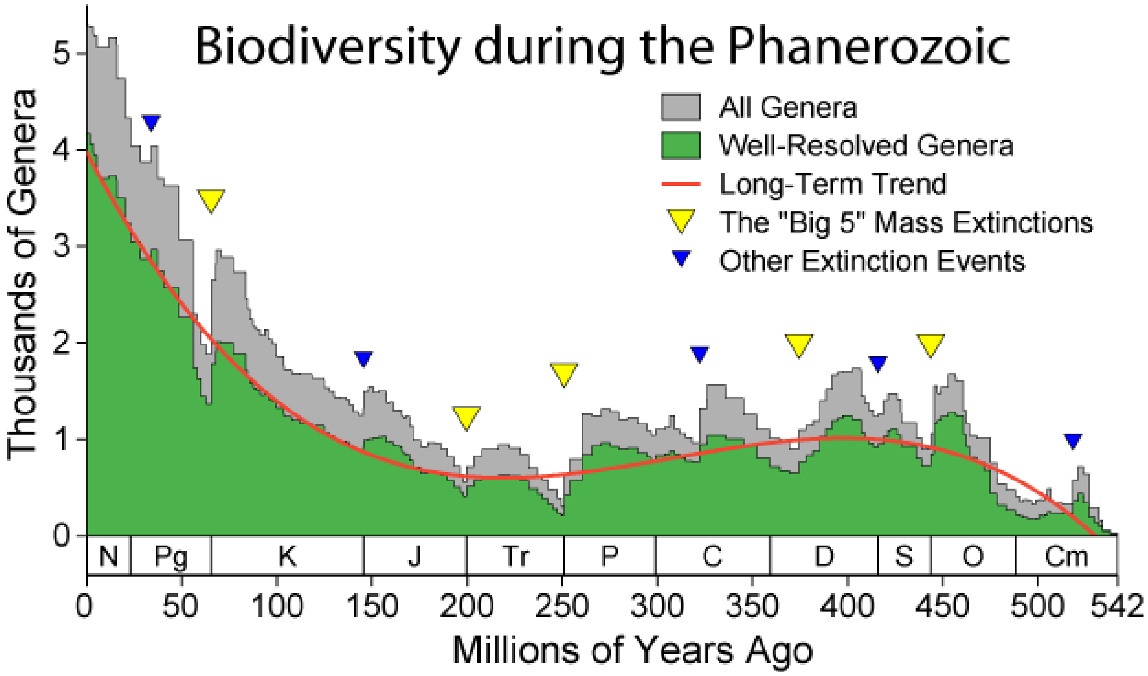
Biodiversity is not evenly distributed. Flora and fauna diversity depends on climate, altitude, soils and the presence of other species. Diversity consistently measures higher in the tropics and in other localized regions such as Cape Floristic Province and lower in polar regions generally. In 2006 many species were formally classified as rare or endangered or threatened; moreover, scientists have estimated that millions more species are at risk which have not been formally recognized. About 40 percent of the 40,177 species assessed using the IUCN Red List criteria are now listed as threatened with extinction—a total of 16,119.

Even though terrestrial biodiversity declines from the equator to the poles, this characteristic is unverified in aquatic ecosystems, especially in marine ecosystems. In addition, several assessments reveal tremendous diversity in higher latitudes. Generally terrestrial biodiversity is up to 25 times greater than ocean biodiversity.

A biodiversity hotspot is a region with a high level of endemic species. Hotspots were first named in 1988 by Dr. Norman Myers. Many hotspots have large nearby human populations. Most hotspots are located in the tropics and most of them are forests.

Brazil's Atlantic Forest is considered one such hotspot, containing roughly 20,000 plant species, 1,350 vertebrates, and millions of insects, about half of which occur nowhere else. The island of Madagascar, particularly the unique Madagascar dry deciduous forests and lowland rainforests, possess a high ratio of endemism. Since the island separated from mainland Africa 65 million years ago, many species and ecosystems have evolved independently. Indonesia's 17,000 islands cover 735,355 square miles (1,904,560 km^2) contain 10% of the world's flowering plants, 12% of mammals and 17% of reptiles, amphibians and birds—along with nearly 240 million people. Many regions of high biodiversity and/or endemism arise from specialized habitats which require unusual adaptations, for example alpine environments in high mountains, or Northern European peat bogs.

Evolution



Apparent marine fossil diversity during the Phanerozoic

Biodiversity is the result of 3.5 billion years of evolution. The origin of life has not been definitely established by science, however some evidence suggests that life may already have been well-established only a few hundred million years after the formation of the Earth. Until approximately 600 million years ago, all life consisted of archaea, bacteria, protozoans and similar single-celled organisms.

The history of biodiversity during the Phanerozoic (the last 540 million years), starts with rapid growth during the Cambrian explosion—a period during which nearly every phylum of multicellular organisms first appeared. Over the next 400 million years or so, global diversity showed little overall trend, but was marked by periodic, massive losses of diversity classified as mass extinction events. A significant loss occurred when rainforests collapsed in the carboniferous. The worst was the Permo-Triassic extinction, 251 million years ago. Vertebrates took 30 million years to recover from this event.

The fossil record suggests that the last few million years featured the greatest biodiversity in history. However, not all scientists support this view, since there is considerable uncertainty as to how strongly the fossil record is biased by the greater availability and preservation of recent geologic sections. Corrected for sampling artifacts, modern biodiversity may not be much different from biodiversity 300 million years ago. Estimates of the present global macroscopic species diversity vary from 2 million to 100 million, with a best estimate of somewhere near 13–14 million, the vast majority arthropods. Diversity appears to increase continually in the absence of natural selection.

Evolutionary diversification

The existence of a "global carrying capacity", limiting the amount of life that can live at once, is debated, as is the question of whether such a limit would also cap the number of species. While records of life in the sea shows a logistic pattern of growth, life on land (insects, plants and tetrapods)shows an exponential rise in diversity. As one author states, "Tetrapods have not yet invaded 64 per cent of potentially habitable modes, and it could be that without human influence the ecological and taxonomic diversity of tetrapods would continue to increase in an exponential fashion until most or all of the available ecospace is filled."

On the other hand, changes through the Phanerozoic correlate much better with the hyperbolic model (widely used in population biology, demography and macrosociology, as well as fossil biodiversity) than with exponential and logistic models. The latter models imply that changes in diversity are guided by a first-order positive feedback (more ancestors, more descendants) and/or a negative feedback arising from resource limitation. Hyperbolic model implies a second-order positive feedback. The hyperbolic pattern of the world population growth arises from a second-order positive feedback between the population size and the rate of technological growth. The hyperbolic character of biodiversity growth can be similarly accounted for by a feedback between diversity and community structure complexity. The similarity between the curves of biodiversity and human population probably comes from the fact that both are derived from the interference of the hyperbolic trend with cyclical and stochastic dynamics.

Most biologists agree however that the period since human emergence is part of a new mass extinction, named the Holocene extinction event, caused primarily by the impact humans are having on the environment. It has been argued that the present rate of extinction is sufficient to eliminate most species on the planet Earth within 100 years.

New species are regularly discovered (on average between 5–10,000 new species each year, most of them insects) and many, though discovered, are not yet classified (estimates are that nearly 90% of all arthropods are not yet classified). Most of the terrestrial diversity is found in tropical forests.

Human benefits

Summer field in Belgium (Hamois). The blue flowers are *Centaurea cyanus* and the red are *Papaver rhoeas*.

Biodiversity supports ecosystem services including air quality, climate (e.g., CO_2 sequestration), water purification, pollination, and prevention of erosion.

Since the stone age, species loss has accelerated above the prior rate, driven by human activity. Estimates of species loss are at a rate 100-10,000 times as fast as is typical in the fossil record.

Non-material benefits include spiritual and aesthetic values, knowledge systems and the value of education.

Agriculture

Amazon Rainforest in Brazil.

The reservoir of genetic traits present in wild varieties and traditionally grown landraces is extremely important in improving crop performance. Important crops, such as potato, banana and coffee, are often derived from only a few genetic strains. Improvements in crop species over the last 250 years have been largely due to incorporating genes from wild varieties and species into cultivars. Crop breeding for beneficial traits has helped to more than double crop production in the last 50 years as a result of the Green Revolution. A biodiverse environment preserves the genome from which such productive genes are drawn.

Crop diversity aids recovery when the dominant cultivar is attacked by a disease or predator:

- The Irish potato blight of 1846 was a major factor in the deaths of one million people and the emigration of another million. It was the result of planting only two potato varieties, both vulnerable to the blight.
- When rice grassy stunt virus struck rice fields from Indonesia to India in the 1970s, 6,273 varieties were tested for resistance. Only one was resistant, an Indian variety, and known to science only since 1966. This variety formed a hybrid with other varieties and is now widely grown.
- Coffee rust attacked coffee plantations in Sri Lanka, Brazil, and Central America in 1970. A resistant variety was found in Ethiopia. Although the diseases are themselves a form of biodiversity.

Monoculture was a contributing factor to several agricultural disasters, including the European wine industry collapse in the late 19th century, and the US Southern Corn Leaf Blight epidemic of 1970.

Higher biodiversity also limits the spread of infectious diseases as many different species act as buffers to them.

Although about 80 percent of humans' food supply comes from just 20 kinds of plants, humans use at least 40,000 species. Many people depend on these species for food, shelter, and clothing. Earth's surviving biodiversity provides resources for increasing the range of food and other products suitable for human use, although the present extinction rate shrinks that potential.

Human health

The diverse forest canopy on Barro Colorado Island, Panama, yielded this display of different fruit

Biodiversity's relevance to human health is becoming an international political issue, as scientific evidence builds on the global health implications of biodiversity loss. This issue is closely linked with the issue of climate change, as many of the anticipated health risks of climate change are associated with changes in biodiversity (e.g. changes in populations and distribution of disease vectors, scarcity of fresh water, impacts on agricultural biodiversity and food resources etc.) Some of the health issues influenced by biodiversity include dietary health and nutrition security, infectious disease, medical science and medicinal resources, social and psychological health. Biodiversity is also known to have an important role in reducing disaster risk, and in post-disaster relief and recovery efforts.

Biodiversity provides critical support for drug discovery and the availability of medicinal resources. A significant proportion of drugs are derived, directly or indirectly, from biological sources; At least 50% of the pharmaceutical compounds on the US market are derived from plants, animals, and micro-organisms, while about 80% of the world population depends on medicines from nature (used in either modern or traditional medical practice) for primary healthcare. Only a tiny fraction of wild species has been investigated for medical potential. Biodiversity has been critical to advances throughout the field of bionics. Evidence from market analysis and biodiversity science indicates that the decline in output from the pharmaceutical sector since the mid-1980s can be attributed to a move away from natural product exploration ("bioprospecting") in favor of genomics and synthetic chemistry; meanwhile, natural products have a long history of supporting significant economic and health innovation. Marine ecosystems are particularly important, although inappropriate bioprospecting can increase biodiversity loss, as well as violating the laws of the communities and states from which the resources are taken.

Business and industry

Agriculture production, pictured is a tractor and a chaser bin

Many industrial materials derive directly from biological sources. These include building materials, fibers, dyes, rubber and oil. Biodiversity is also important to the security of resources such as water, timber, paper, fiber, and food. As a result, biodiversity loss is a significant risk factor in business development and a threat to long term economic sustainability.

Leisure, cultural and aesthetic value

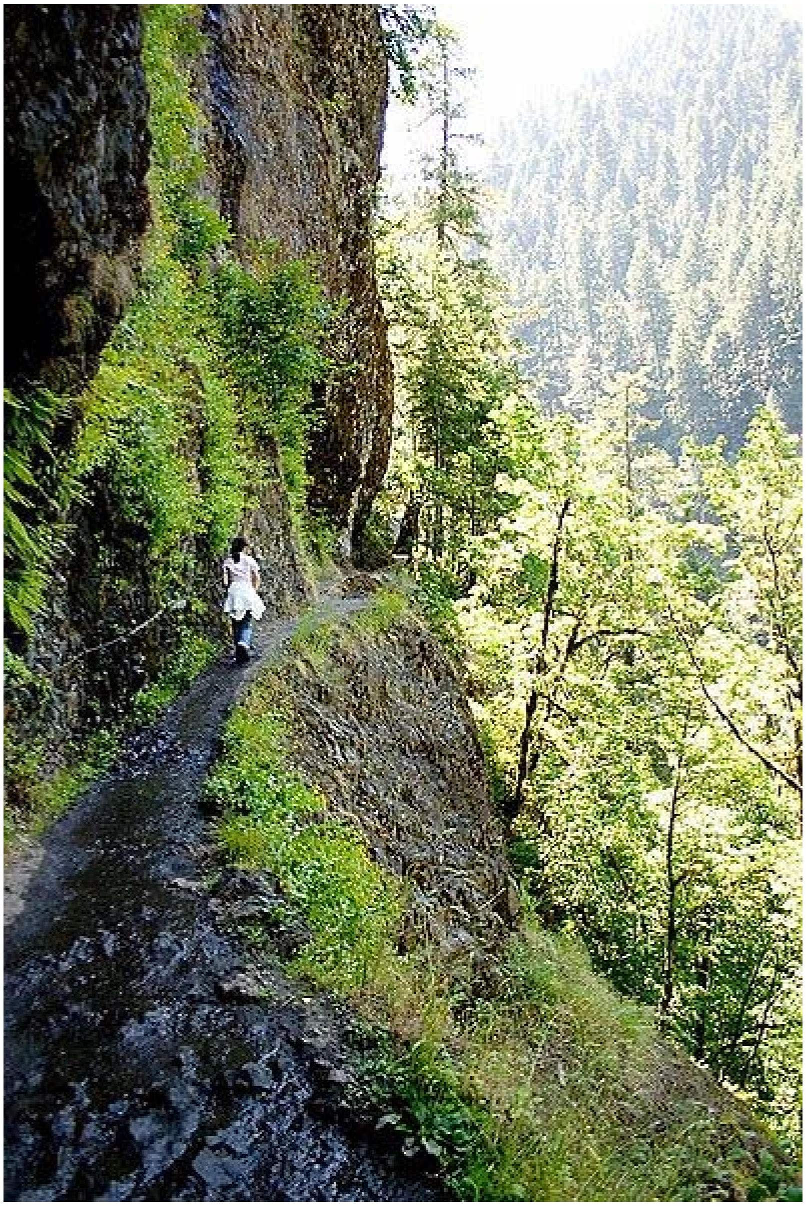
Biodiversity enriches leisure activities such as hiking, birdwatching or natural history study. Biodiversity inspires musicians, painters, sculptors, writers and other artists. Many cultures view themselves as an integral part of the natural world which requires them to respect other living organisms.

Popular activities such as gardening, fishkeeping and specimen collecting strongly depend on biodiversity. The number of species involved in such pursuits is in the tens of thousands, though the majority do not enter commerce.

The relationships between the original natural areas of these often exotic animals and plants and commercial collectors, suppliers, breeders, propagators and those who promote their understanding and enjoyment are complex and poorly understood. The general public responds well to exposure to rare and unusual organisms, reflecting their inherent value.

Philosophically it could be argued that biodiversity has intrinsic aesthetic and spiritual value to mankind *in and of itself*. This idea can be used as a counterweight to the notion that tropical forests and other ecological realms are only worthy of conservation because of the services they provide.

Other services



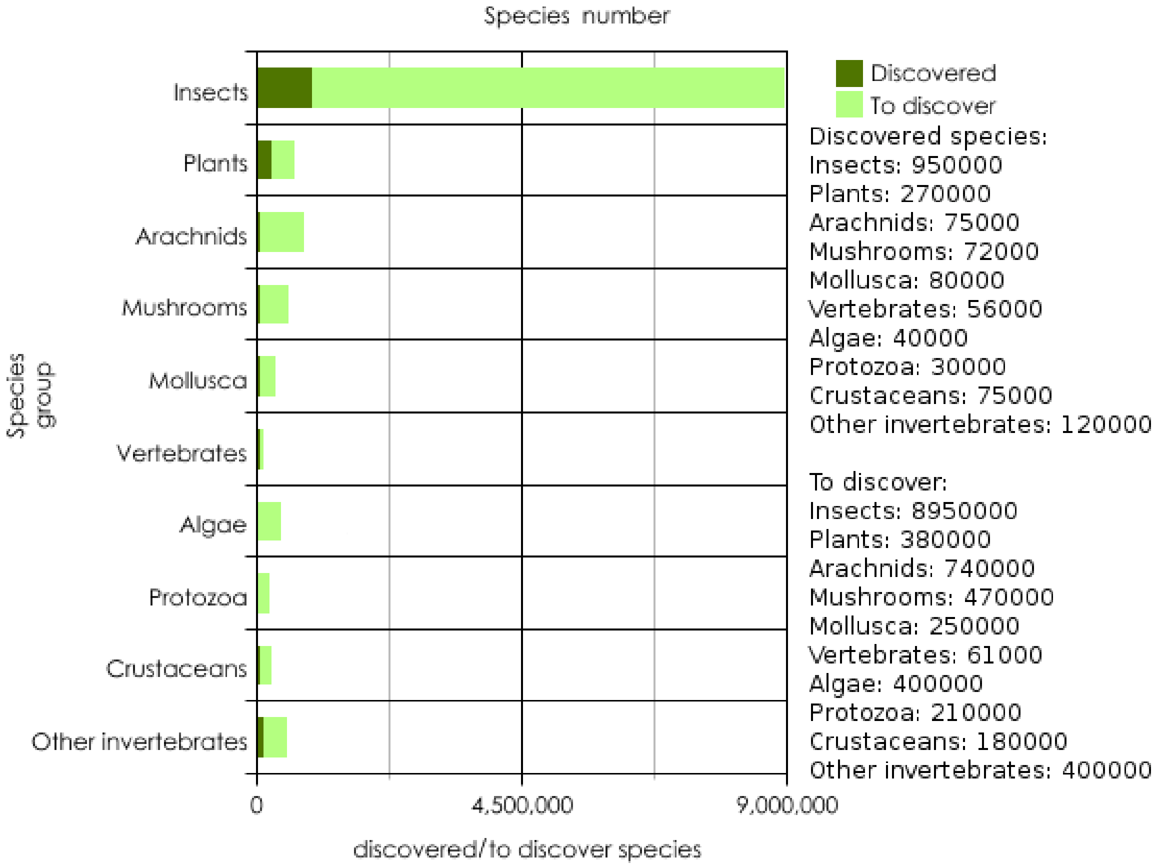
Eagle Creek, Oregon hiking

Biodiversity supports many ecosystem services that are often not readily visible. It plays a part in regulating the chemistry of our atmosphere and water supply. Biodiversity is directly involved in water purification, recycling nutrients and providing fertile soils. Experiments with controlled environments have shown that humans cannot easily build ecosystems to support human needs; for example insect pollination cannot be mimicked, and that activity alone represents tens of billions of dollars in ecosystem services per year to humankind.

Ecosystem stability is also positively related to biodiversity, protecting against disruption by extreme weather or human exploitation.

Polar bears on the sea ice of the Arctic Ocean, near the North Pole.

Number of species



Undiscovered and discovered species

According to the Global Taxonomy Initiative and the European Distributed Institute of Taxonomy, the *total* number of species for some phyla may be much higher than what was known in 2010:

- 10–30 million insects; (of some 0.9 million we know today)
- 5–10 million bacteria;
- 1.5 million fungi; (of some 0.4 million we know today)
- 1 million mites
- The number of microbial species is not reliably known, but the Global Ocean Sampling Expedition dramatically increased the estimates of genetic diversity by identifying an enormous number of new genes from near-surface plankton samples at various marine locations, initially over the 2004-2006 period. The findings may eventually cause a significant change in the way science defines species and other taxonomic categories.

Since the rate of extinction has increased, many extant species may become extinct before they are described.

Species loss rates

During the last century, decreases in biodiversity have been increasingly observed. In 2007, German Federal Environment Minister Sigmar Gabriel cited estimates that up to 30% of all species will be extinct by 2050. Of these, about one eighth of known plant species are threatened with extinction. Estimates reach as high as 140,000 species per year (based on Species-area theory). This figure indicates unsustainable ecological practices, because few species emerge each year. Almost all scientists acknowledge that the rate of species loss is greater now than at any time in human history, with extinctions occurring at rates hundreds of times higher than background extinction rates.

Threats

Jared Diamond describes an "Evil Quartet" of habitat destruction, overkill, introduced species, and secondary extinctions. Edward O. Wilson prefers the acronym HIPPO, standing for habitat destruction, invasive species, pollution, human over population, and over-harvesting. The most authoritative classification in use today is IUCN’s Classification of Direct Threats which has been adopted by major international conservation organizations such as the US Nature Conservancy, the World Wildlife Fund, Conservation International, and Birdlife International.

Habitat destruction

Deforestation and increased road-building in the Amazon Rainforest are a significant concern because of increased human encroachment upon wild areas, increased resource extraction and further threats to biodiversity.

Habitat destruction has played a key role in extinctions, especially related to tropical forest destruction. Factors contributing to habitat loss are: overpopulation, deforestation, pollution (air pollution, water pollution, soil contamination) and global warming or climate change.

Habitat size and numbers of species are systematically related. Physically larger species and those living at lower latitudes or in forests or oceans are more sensitive to reduction in habitat area. Conversion to "trivial" standardized ecosystems (e.g., monoculture following deforestation) effectively destroys habitat for the more diverse species that preceded the conversion. In some countries lack of property rights or lax law/regulatory enforcement necessarily leads to biodiversity loss (degradation costs having to be supported by the community).

A 2007 study conducted by the National Science Foundation found that biodiversity and genetic diversity are codependent—that diversity among species requires diversity within a species, and vice versa. "If any one type is removed from the system, the cycle can break down, and the community becomes dominated by a single species." At present, the most threatened ecosystems are found in fresh water, according to the Millennium Ecosystem Assessment 2005, which was confirmed by the "Freshwater Animal Diversity Assessment", organised by the biodiversity platform, and the French Institut de recherche pour le développement (MNHNP).

Co-extinctions are a form of habitat destruction. Co-extinction occurs when the extinction or decline in one accompanies the other, such as in plants and beetles.

Introduced and invasive species

Male *Lophura nycthemera* (Silver Pheasant), a native of East Asia that has been introduced into parts of Europe for ornamental reasons

Barriers such as large rivers, seas, oceans, mountains and deserts encourage diversity by enabling independent evolution on either side of the barrier. Invasive species occur when those barriers are blurred. Without barriers such species occupy new niches, substantially reducing diversity. Repeatedly humans have helped these species circumvent these barriers, introducing them for food and other purposes. This has occurred on a time scale much shorter than the eons that historically have been required for a species to extend its range.

Not all introduced species are invasive, nor all invasive species deliberately introduced. In cases such as the zebra mussel, invasion of US waterways was unintentional. In other cases, such as mongooses in Hawaii, the introduction is deliberate but ineffective (nocturnal rats were not vulnerable to the diurnal mongoose!). In other cases, such as oil palms in Indonesia and Malaysia, the introduction produces substantial economic benefits, but the benefits are accompanied by costly unintended consequences.

Finally, an introduced species may unintentionally injure a species that depends on the species it replaces. In Belgium, Prunus spinosa from Eastern Europe leafs much sooner than its West European counterparts, disrupting the feeding habits of the Thecla betulae butterfly (which feeds on the leaves). Introducing new species often leaves endemic and other local species unable to compete with the exotic species and unable to survive. The exotic organisms may be predators, parasites, or may simply outcompete indigenous species for nutrients, water and light.

At present, several countries have already imported so many exotic species, particularly agricultural and ornamental plants, that the own indigenous fauna/flora may be outnumbered.

Genetic pollution

Endemic species can be threatened with extinction through the process of genetic pollution, i.e. uncontrolled hybridization, introgression and genetic swamping. Genetic pollution leads to homogenization or replacement of local genomes as a result of either a numerical and/or fitness advantage of an introduced species. Hybridization and introgression are side-effects of introduction and invasion. These phenomena can be especially detrimental to rare species that come into contact with more abundant ones. The abundant species can interbreed with the rare species, swamping its gene pool. This problem is not always apparent from morphological (outward appearance) observations alone. Some degree of gene flowis normal adaptation, and not all gene and genotype constellations can be preserved. However, hybridization with or without introgression may, nevertheless, threaten a rare species' existence.

Overexploitation

Overexploitation occurs when a resource is consumed at an unsustainable rate. This occurs on land in the form of overhunting, excessive logging, poor soil conservation in agriculture and the illegal wildlife trade. Joe Walston, director of the Wildlife Conservation Society’s Asian programs, called the latter the "single largest threat" to biodiversity in Asia. The international trade of endangered species is second in size only to drug trafficking.

About 25% of world fisheries are now overfished to the point where their current biomass is less than the level that maximizes their sustainable yield.

The overkill hypothesis explains why earlier megafaunal extinctions occurred within a relatively short period of time. This can be connected with human migration.

Hybridization, genetic pollution/erosion and food security

The Yecoro wheat (right) cultivar is sensitive to salinity, plants resulting from a hybrid cross with cultivar W4910 (left) show greater tolerance to high salinity

In agriculture and animal husbandry, the Green Revolution popularized the use of conventional hybridization to increase yield. Often hybridized breeds originated in developed countries and were further hybridized with local varieties in the developing world to create high yield strains resistant to local climate and diseases. Local governments and industry have been pushing hybridization. Formerly huge gene pools of various wild and indigenous breeds have collapsed causing widespread genetic erosion and genetic pollution. This has resulted in loss of genetic diversity and biodiversity as a whole.

(GM organisms) have genetic material altered by genetic engineering procedures such as recombinant DNA technology. GM crops have become a common source for genetic pollution, not only of wild varieties but also of domesticated varieties derived from classical hybridization.

Genetic erosion coupled with genetic pollution may be destroying unique genotypes, thereby creating a hidden crisis which could result in a severe threat to our food security. Diverse genetic material could cease to exist which would impact our ability to further hybridize food crops and livestock against more resistant diseases and climatic changes.

Climate Change

Global warming is also considered to be a major threat to global biodiversity. For example coral reefs -which are biodiversity hotspots- will be lost in 20 to 40 years if global warming continues at the current trend.

In 2004, an international collaborative study on four continents estimated that 10 percent of species would become extinct by 2050 because of global warming. "We need to limit climate change or we wind up with a lot of species in trouble, possibly extinct," said Dr. Lee Hannah, a co-author of the paper and chief climate change biologist at the Center for Applied Biodiversity Science at Conservation International.

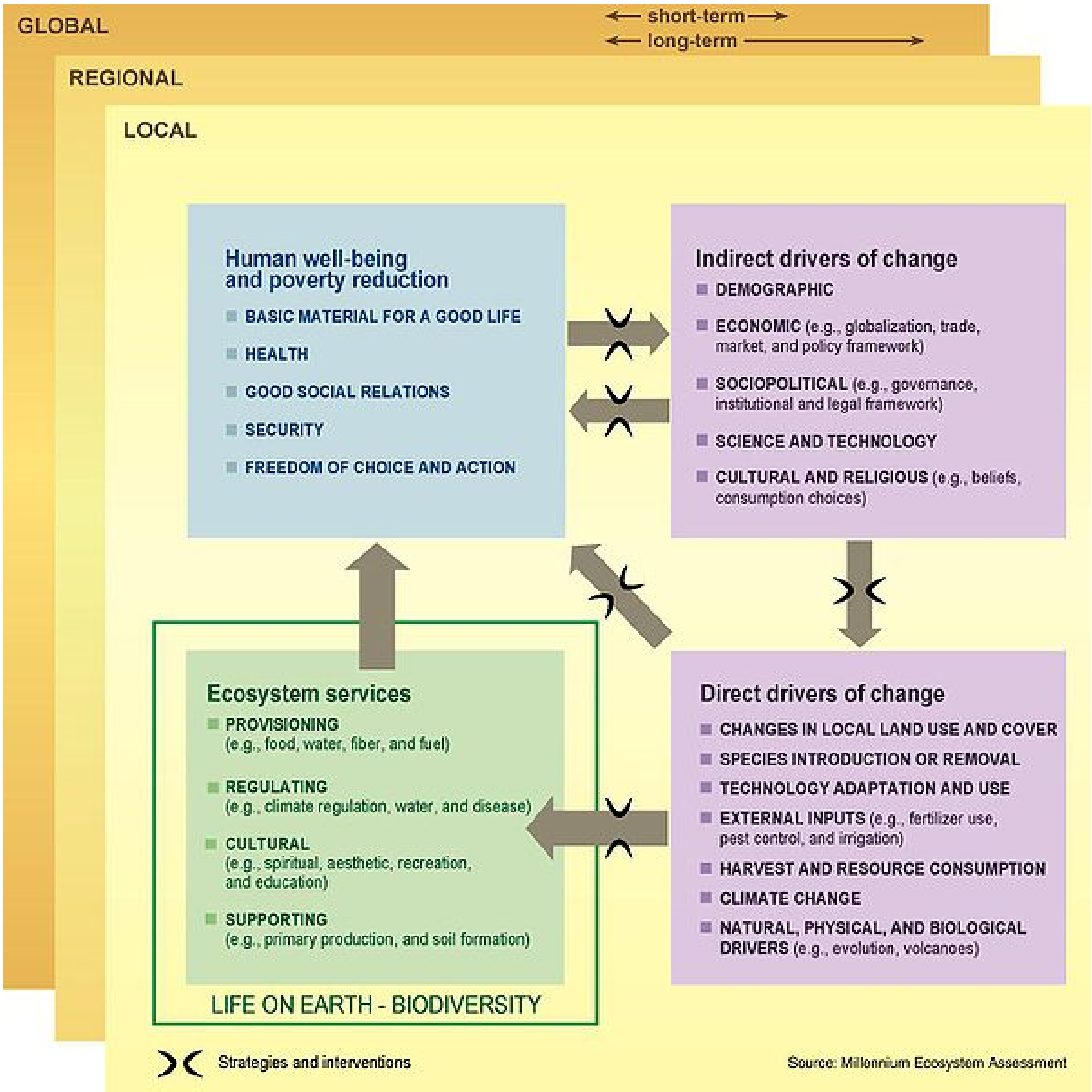
Overpopulation

From 1950 to 2005, world population increased from 2.5 billion to 6.5 billion and is forecast to reach a plateau of more than 9 billion during the 21st century. Sir David King, former chief scientific adviser to the UK government, told a parliamentary inquiry: "It is self-evident that the massive growth in the human population through the 20th century has had more impact on biodiversity than any other single factor."

The Holocene extinction

Rates of decline in biodiversity in this sixth mass extinction match or exceed rates of loss in the five previous mass extinction events in the fossil record. Loss of biodiversity results in the loss of natural capital that supplies ecosystem goods and services. The economic value of 17 ecosystem services for Earth's biosphere (calculated in 1997) has an estimated value of US$ 33 trillion (3.3×10^{13}) per year.

Conservation



A schematic image illustrating the relationship between biodiversity, ecosystem services, human well-being, and poverty. The illustration shows where conservation action, strategies and plans can influence the drivers of the current biodiversity crisis at local, regional, to global scales.

The retreat of Aletsch Glacier in the Swiss Alps (situation in 1979, 1991 and 2002), due to global warming.

Conservation biology matured in the mid-20th century as ecologists, naturalists, and other scientists began to research and address issues pertaining to global biodiversity declines.

The conservation ethic advocates management of natural resources for the purpose of sustaining biodiversity in species, ecosystems, the evolutionary process, and human culture and society.

Conservation biology is reforming around strategic plans to protect biodiversity. Preserving global biodiversity is a priority in strategic conservation plans that are designed to engage public policy and concerns affecting local, regional and global scales of communities, ecosystems, and cultures. Action plans identify ways of sustaining human well-being, employing natural capital, market capital, and ecosystem services.

Protection and restoration techniques

The most powerful technique is to preserve habitat.

Exotic species removal allows less competitive species to recover their ecological niches. Exotic species that have become a pest can be identified taxonomically (e.g. with Digital Automated Identification SYstem (DAISY), using the barcode of life. Removal is practical only given large groups of individuals due to the econimic cost.

Once the preservation of the remaining native species in an area is assured. "missing" species can be identified and reintroduced using databases such as the *Encyclopedia of Life* and the Global Biodiversity Information Facility.

Other techniques include:

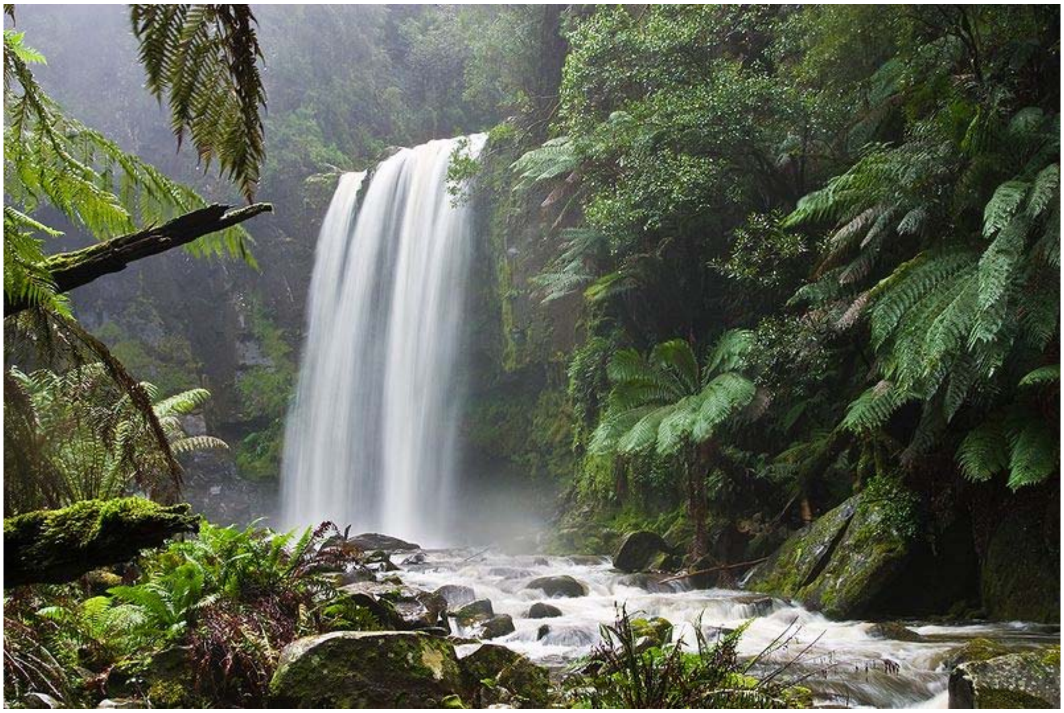
- Biodiversity banking places a monetary value on biodiversity. One example is the Australian Native Vegetation Management Framework.
- Gene banks are collections of specimens and genetic material. Some banks intend to reintroduce banked species to the ecosystem (e.g. via tree nurseries).
- Reducing and better targeting of pesticides allows more species to survive in agricultural and urbanized areas.
- Location-specific approaches are less useful for protecting migratory species. One approach is to create wildlife corridors that correspond to the animals' movements. National and other boundaries can complicate corridor creation.

Resource allocation

Focusing on limited areas of higher potential biodiversity promises greater immediate return on investment than spreading resources evenly or focusing on areas of little diversity but greater interest in biodiversity.

A second strategy focuses on areas that retain most of their original diversity, which typically require little or no restoration. These are typically non-urbanized, non-agricultural areas. Tropical areas often fit both criteria, given their natively high diversity and relative lack of development.

Legal status



A great deal of work is occurring to preserve the natural characteristics of Hopetoun Falls, Australia while continuing to allow visitor access.

Biodiversity is taken into account in some political and judicial decisions:

- The relationship between law and ecosystems is very ancient and has consequences for biodiversity. It is related to private and public property rights. It can define protection for threatened ecosystems, but also some rights and duties (for example, fishing and hunting rights).
- Law regarding species is more recent. It defines species that must be protected because they may be threatened by extinction. The U.S. Endangered Species Act is an example of an attempt to address the "law and species" issue.
- Laws regarding gene pools are only about a century old. Domestication and plant breeding methods are not new, but advances in genetic engineering has led to tighter laws covering distribution of genetically modified organisms, gene patents and process patents. Governments struggle to decide whether to focus on for example, genes, genomes, or organisms and species.

Global agreements such as the Convention on Biological Diversity), give **sovereign national rights over biological resources** (not property). The agreements commit countries to **conserve biodiversity**, **develop resources for sustainability** and **share the benefits** resulting from their use. Biodiverse countries that allow bioprospecting or collection of natural products, expect a share of the benefits rather than allowing the individual or institution that discovers/exploits the resource to capture them privately. Bioprospecting can become a type of biopiracy when such principles are not respected.

Sovereignty principles can rely upon what is better known as Access and Benefit Sharing Agreements (ABAs). The Convention on Biodiversity implies informed consent between the source country and the collector, to establish which resource will be used and for what, and to settle on a fair agreement on benefit sharing.

Uniform approval for use of biodiversity as a legal standard has not been achieved, however. Bosselman argues that biodiversity should not be used as a legal standard, claiming that the remaining areas of scientific uncertainty cause unacceptable administrative waste and increase litigation without promoting preservation goals.

Analytical limits

Taxonomic and size relationships

Less than 1% of all species that have been described have been studied beyond simply noting their existence. The vast majority of Earth's species are microbial. Contemporary biodiversity physics is "firmly fixated on the visible [macroscopic] world". For example, microbial life is metabolically and environmentally more diverse than multicellular life. "On the tree of life, based on analyses of small-subunit ribosomal RNA, visible life consists of barely noticeable twigs. The inverse relationship of size and population recurs higher on the evolutionary ladder—"to a first approximation, all multicellular species on Earth are insects". Insect extinction rates are high—supporting the Holocene extinction hypothesis.