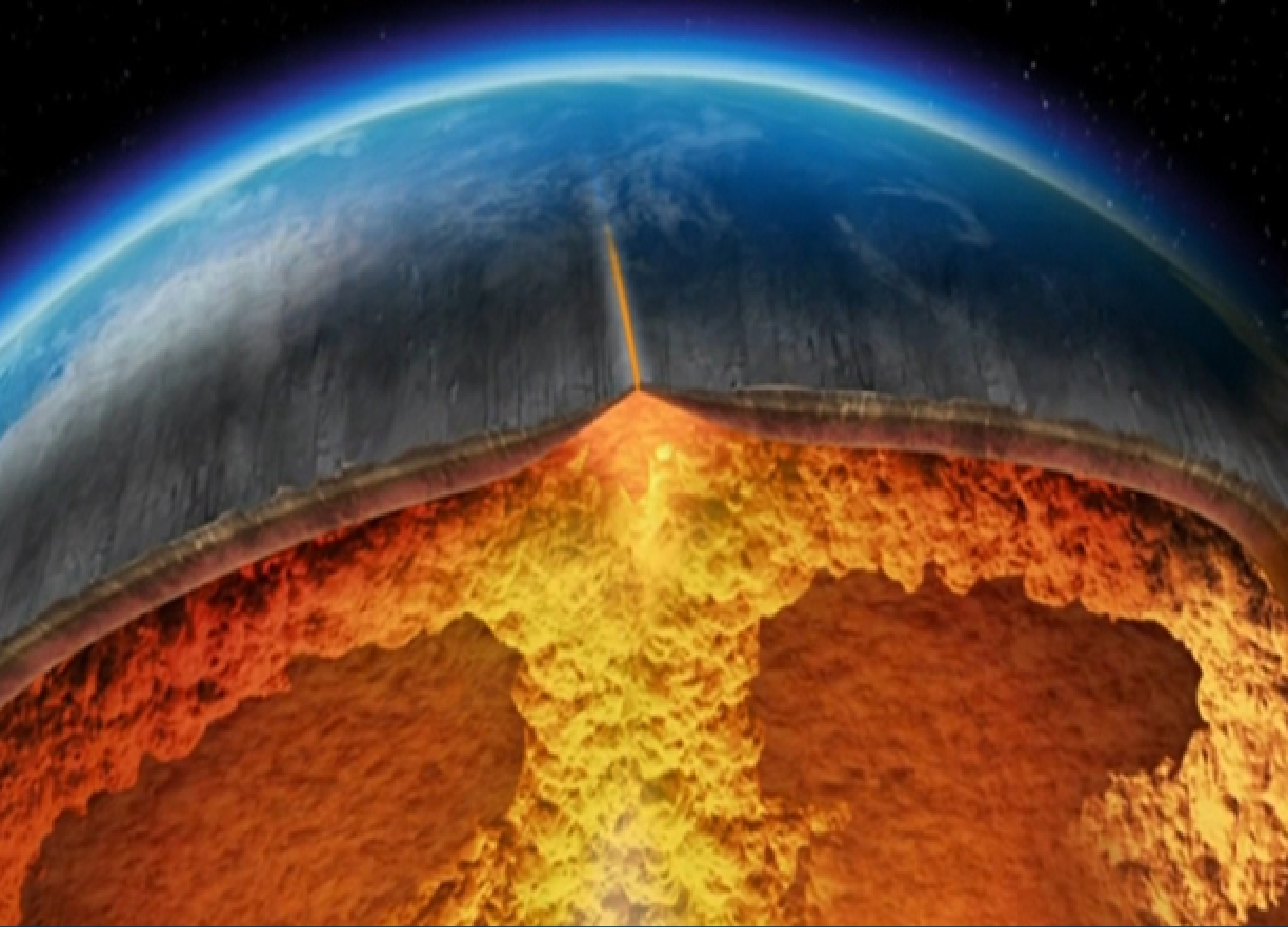


History of the Earth, Solar System and Human Evolution

Alfonso Mireles

First Edition, 2012

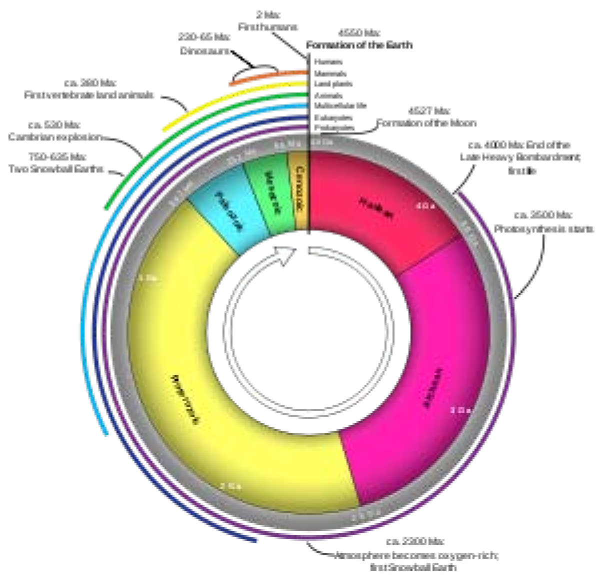
ISBN 978-81-323-1570-4

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

Table of Contents

Chapter- 1

History of the Earth



Geological time put in a diagram called a geological clock, showing the relative lengths of the eons of the Earth's history

The **history of the Earth** describes the most important events and fundamental stages in the development of the planet Earth from its formation 4.54 billion years ago to the present day. Nearly all branches of natural science have contributed to the understanding of the main events of the Earth's past. The age of Earth is approximately one-third of the age of the universe. Immense geological and biological changes have occurred during that time span.

4.54 Ga: Hadean and Archaean eons

Starting with the Earth's formation by accretion from the solar nebula 4.54 billion years ago (4.54 Ga), the first eon in the Earth's history is called the Hadean. It lasted until the Archaean eon, which began 3.8 Ga. The oldest rocks found on Earth date to about 4.0 Ga, and the oldest detrital zircon crystals in some rocks have been dated to about 4.4 Ga, close to the formation of the Earth's crust and the Earth itself. Because not much material from this time is preserved, little is known about Hadean times, but scientists hypothesize at an estimated 4.53 Ga, shortly after formation of an initial crust, the proto-Earth was impacted by a smaller protoplanet, which ejected part of the mantle and crust into space and created the Moon.

During the Hadean, the Earth's surface was under a continuous bombardment by meteorites, and volcanism must have been severe due to the large heat flow and geothermal gradient. The detrital zircon crystals dated to 4.4 Ga show evidence of having undergone contact with liquid water, suggesting that the planet already had oceans or seas at that time. From crater counts on other celestial bodies it is inferred that a period of intense meteorite impacts, called the "Late Heavy Bombardment", began about 4.1 Ga, and concluded around 3.8 Ga, at the end of the Hadean.

By the beginning of the Archaean, the Earth had cooled significantly. It would have been impossible for most present day life forms to exist due to the composition of the Archaean atmosphere, which lacked oxygen and an ozone layer. Nevertheless it is believed that primordial life began to evolve by the early Archaean, with some possible fossil finds dated to around 3.5 Ga. Some researchers, however, speculate that life could have begun during the early Hadean, as far back as 4.4 Ga, surviving the possible Late Heavy Bombardment period in hydrothermal vents below the Earth's surface.

4.6 Ga: Solar System

An artist's impression of protoplanetary disk

The Solar System (including the Earth) formed from a large, rotating cloud of interstellar dust and gas called the solar nebula, orbiting the Milky Way's galactic center. It was composed of hydrogen and helium created shortly after the Big Bang 13.7 Ga and heavier elements ejected by supernovas. About 4.6 Ga, the solar nebula began to contract, possibly due to the shock wave of a nearby supernova. Such a shock wave would have also caused the nebula to rotate and gain angular momentum. As the cloud began to accelerate its rotation, gravity and inertia flattened it into a protoplanetary disk oriented perpendicularly to its axis of rotation. Most of the mass concentrated in the middle and began to heat up, but small perturbations due to collisions and the angular momentum of other large debris created the means by which protoplanets up to several kilometres in length began to form, orbiting the nebular center.

The infall of material, increase in rotational speed and the crush of gravity created an enormous amount of kinetic energy at the center. Its inability to transfer that energy away through any other process at a rate capable of relieving the build-up resulted in the disk's center heating up. Ultimately, nuclear fusion of hydrogen into helium began, and eventually, after contraction, a T Tauri star ignited to create the Sun. Meanwhile, as gravity caused matter to condense around the previously perturbed objects outside the gravitational grasp of the new sun, dust particles and the rest of the protoplanetary disk began separating into rings. Successively larger fragments collided with one another and became larger objects, ultimately becoming protoplanets. These included one collection about 150 million kilometers from the center: Earth. The planet formed about 4.54 billion years ago (within an uncertainty of 1%) and was largely completed within 10–20 million years. The solar wind of the newly formed T Tauri star cleared out most of the material in the disk that had not already condensed into larger bodies.

Computer simulations have shown that planets with distances equal to the terrestrial planets in our solar system can be created from a protoplanetary disk. The now widely accepted nebular hypothesis suggests that the same process, which gave rise to the solar system's planets, produces accretion disks around virtually all newly forming stars in the universe, some of which yield planets.

Origin of the Earth's core and first atmosphere

The Proto-Earth grew by accretion, until the inner part of the protoplanet was hot enough to melt the heavy, siderophile metals. Such liquid metals, with now higher densities, began to sink to the Earth's center of mass. This so called iron catastrophe resulted in the separation of a primitive mantle and a (metallic) core only 10 million years after the Earth began to form, producing the layered structure of Earth and setting up the formation of Earth's magnetic field.

During the accretion of material to the protoplanet, a cloud of gaseous silica must have surrounded the Earth, to condense afterwards as solid rocks on the surface. What was left surrounding the planet was an early atmosphere of light (atmophile) elements from the solar nebula, mostly hydrogen and helium, but the solar wind and Earth's heat would have driven off this atmosphere.

This changed when Earth accreted to about 40% its present radius, and gravitational attraction retained an atmosphere which included water.

4.52 Ga: The giant impact hypothesis

The Earth's relatively large natural satellite, the Moon, is unique. During the Apollo program, rocks from the Moon's surface were brought to Earth. Radiometric dating of these rocks has shown the Moon to be 4527 ± 10 million years old, about 30 to 55 million years younger than other bodies in the solar system. (New evidence suggests the Moon formed even later, 4.48±0.02 Ga, or 70–110 Ma after the start of the Solar System.) Another notable feature is the relatively low density of the Moon, which must mean it does not have a large metallic core, like all other terrestrial bodies in the solar system. The Moon has a bulk composition closely resembling the Earth's mantle and crust together, without the Earth's core. This has led to the giant impact hypothesis, the idea that the Moon was formed during a giant impact of the proto-Earth with another protoplanet by accretion of the material blown off the mantles of the proto-Earth and impactor.

The impactor, sometimes named Theia, is thought to have been a little smaller than the current planet Mars. It could have formed by accretion of matter about 150 million kilometres from the Sun and Earth, at their fourth or fifth Lagrangian point. Its orbit may have been stable at first, but destabilized as Theia's mass increased due to the accretion of matter. Theia oscillated in larger and larger orbits around the Lagrangian point until it finally collided with Earth about 4.533 Ga. Models reveal that when an impactor this size struck the proto-Earth at a low angle and relatively low speed (8–20 km/sec), much material from the mantles and crusts of the proto-Earth and the impactor was ejected into space, where much of it stayed in orbit around the Earth. This material would eventually form the Moon. However, the metallic cores of the impactor would have sunk through the Earth's mantle to fuse with the Earth's core, depleting the Moon of metallic material. The giant impact hypothesis thus explains the Moon's abnormal composition. The ejecta in orbit around the Earth could have condensed into a single body within a couple of weeks. Under the influence of its own gravity, the ejected material became a more spherical body: the Moon.

The radiometric ages show the Earth existed already for at least 10 million years before the impact, enough time to allow for differentiation of the Earth's primitive mantle and core. Then, when the impact occurred, only material from the mantle was ejected, leaving the Earth's core of heavy siderophile elements untouched.

The impact had some important consequences for the young Earth. It released an enormous amount of energy, causing both the Earth and Moon to be completely molten. Immediately after the impact, the Earth's mantle was vigorously convecting, the surface was a large magma ocean. The planet's first atmosphere must have been completely blown away by the impact. The impact is also thought to have changed Earth's axis to produce the large 23.5° axial tilt that is responsible for Earth's seasons (a simple, ideal model of the planets' origins would have axial tilts of 0° with no recognizable seasons). It may also have sped up Earth's rotation.

4.3 Ga: Oceans and atmosphere

Because the Earth lacked an atmosphere immediately after the giant impact, cooling must have occurred quickly. Within 150 million years, a solid crust with a basaltic composition must have formed. The felsic continental crust of today did not yet exist. Within the Earth, further differentiation could only begin when the mantle had at least partly solidified again. Nevertheless, during the early Archaean (about 3.0 Ga) the mantle was still much hotter than today, probably around 1600°C. This means the fraction of partially molten material was still much larger than today.

Steam escaped from the crust, and more gases were released by volcanoes, completing the second atmosphere. Additional water was imported by bolide collisions, probably from asteroids ejected from the outer asteroid belt under the influence of Jupiter's gravity.

The large amount of water on Earth can never have been produced by volcanism and degassing alone. It is assumed the water was derived from impacting comets that contained ice. Though most comets are today in orbits farther away from the Sun than Neptune, computer simulations show they were originally far more common in the inner parts of the solar system. However, most of the water on Earth was probably derived from small impacting protoplanets, objects comparable with today's small icy moons of the outer planets. Impacts of these objects could have enriched the terrestrial planets (Mercury, Venus, the Earth and Mars) with water, carbon dioxide, methane, ammonia, nitrogen and other volatiles. If all water on Earth was derived from comets alone, millions of comet impacts would be required to support this theory. Computer simulations illustrate that this is not an unreasonable number.

As the planet cooled, clouds formed. Rain created the oceans. Recent evidence suggests the oceans may have begun forming by 4.2 Ga, or as early as 4.4 Ga. In any event, by the start of the Archaean eon the Earth was already covered with oceans. The new atmosphere probably contained water vapor, carbon dioxide, nitrogen, and smaller amounts of other gases. As the output of the Sun was only 70% of the current amount, significant amounts of greenhouse gas in the atmosphere most likely prevented the surface water from freezing. Free oxygen would have been bound by hydrogen or minerals on the surface. Volcanic activity was intense and, without an ozone layer to hinder its entry, ultraviolet radiation flooded the surface.

Lithified stromatolites on the shores of Lake Thetis (Western Australia). Stromatolites are formed by colonies of single celled organisms like cyanobacteria or chlorophyta. These colonies of algae entrap sedimentary grains, thus forming the draped sedimentary layers of a stromatolite. Archaean stromatolites are the first direct fossil traces of life on Earth, even though little preserved fossilized cells have been found inside them. The Archaean and Proterozoic oceans could have been full of algal mats like these.

4.0 Ga: The first continents

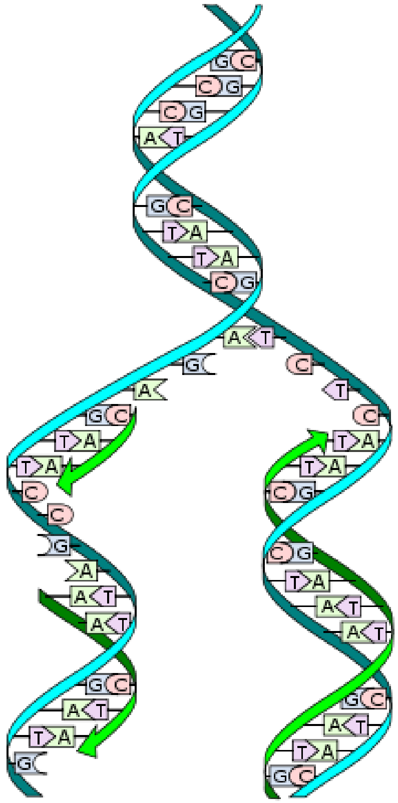
Mantle convection, the process that drives plate tectonics today, is a result of heat flow from the core to the Earth's surface. It involves the creation of rigid tectonic plates at mid-oceanic ridges. These plates are destroyed by subduction into the mantle at subduction zones. The inner Earth was warmer during the Hadean and Archaean eons, so convection in the mantle must have been faster. When a process similar to present day plate tectonics did occur, this would have gone faster too. Most geologists believe that during the Hadean and Archaean, subduction zones were more common, and therefore tectonic plates were smaller.

The initial crust, formed when the Earth's surface first solidified, totally disappeared from a combination of this fast Hadean plate tectonics and the intense impacts of the Late Heavy Bombardment. It is, however, assumed that this crust must have been basaltic in composition, like today's oceanic crust, because little crustal differentiation had yet taken place. The first larger pieces of continental crust, which is a product of differentiation of lighter elements during partial melting in the lower crust, appeared at the end of the Hadean, about 4.0 Ga. What is left of these first small continents are called cratons. These pieces of late Hadean and early Archaean crust form the cores around which today's continents grew.

The oldest rocks on Earth are found in the North American craton of Canada. They are tonalites from about 4.0 Ga. They show traces of metamorphism by high temperature, but also sedimentary grains that have been rounded by erosion during transport by water, showing rivers and seas existed then.

Cratons consist primarily of two alternating types of terranes. The first are so called greenstone belts, consisting of low grade metamorphosed sedimentary rocks. These "greenstones" are similar to the sediments today found in oceanic trenches, above subduction zones. For this reason, greenstones are sometimes seen as evidence for subduction during the Archaean. The second type is a complex of felsic magmatic rocks. These rocks are mostly tonalite, trondhjemite or granodiorite, types of rock similar in composition to granite (hence such terranes are called TTG-terranes). TTG-complexes are seen as the relicts of the first continental crust, formed by partial melting in basalt. The alternation between greenstone belts and TTG-complexes is interpreted as a tectonic situation in which small proto-continents were separated by a thorough network of subduction zones.

3.5 Ga: Life



The replicator in virtually all known life is deoxyribonucleic acid. DNA is far more complex than the original replicator and its replication systems are highly elaborate.

The details of the origin of life are unknown, but the basic principles have been established. There are two schools of thought about the origin of life. One suggests that organic components arrived on Earth from space, while the other argues that they originated on Earth. Nevertheless, both schools suggest similar mechanisms by which life initially arose.

If life arose on Earth, the timing of this event is highly speculative—perhaps it arose around 4 Ga. It is possible that, as a result of repeated formation and destruction of oceans during that time period caused by high energy asteroid bombardment, life may have arisen and been extinguished more than once.

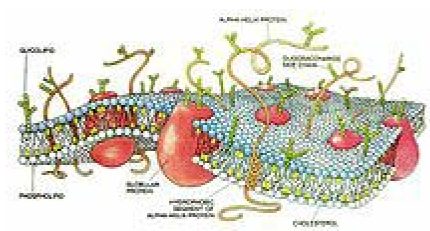
In the energetic chemistry of early Earth, a molecule gained the ability to make copies of itself — a replicator. (More accurately, it promoted the chemical reactions which produced a copy of itself.) The replication was not always accurate: some copies were slightly different from their parent.

If the change destroyed the copying ability of the molecule, the molecule did not produce any copies, and the line “died out”. On the other hand, a few rare changes might have made the molecule replicate faster or better: those “strains” would become more numerous and “successful”. This is an early example of evolution on abiotic material. The variations present in matter and molecules combined with the universal tendency for systems to move towards a lower energy state allowed for an early method of natural selection. As choice raw materials (“food”) became depleted, strains which could utilize different materials, or perhaps halt the development of other strains and steal their resources, became more numerous.

The nature of the first replicator is unknown because its function was long since superseded by life’s current replicator, DNA. Several models have been proposed explaining how a replicator might have developed. Different replicators have been posited, including organic chemicals such as modern proteins, nucleic acids, phospholipids, crystals, or even quantum systems. There is currently no way to determine whether any of these models closely fits the origin of life on Earth.

One of the older theories, one which has been worked out in some detail, will serve as an example of how this might occur. The high energy from volcanoes, lightning, and ultraviolet radiation could help drive chemical reactions producing more complex molecules from simple compounds such as methane and ammonia. Among these were many of the simpler organic compounds, including nucleobases and amino acids, which are the building blocks of life. As the amount and concentration of this “organic soup” increased, different molecules reacted with one another. Sometimes more complex molecules would result—perhaps clay provided a framework to collect and concentrate organic material.

Certain molecules could speed up a chemical reaction. All this continued for a long time, with reactions occurring at random, until by chance it produced a replicator molecule. In any case, at some point, the function of the replicator was superseded by DNA; all known life (except some viruses and prions) use DNA as their replicator, in an almost identical manner.



A small section of a cell membrane. This modern cell membrane is far more sophisticated than the original simple phospholipid bilayer (the small blue spheres with two tails). Proteins and carbohydrates serve various functions in regulating the passage of material through the membrane and in reacting to the environment.

Modern life has its replicating material packaged inside a cellular membrane. It is easier to understand the origin of the cell membrane than the origin of the replicator, because a cell membrane is made of phospholipid molecules, which often form a bilayer spontaneously when placed in water. Under certain conditions, many such spheres can be formed.

The prevailing theory is that the membrane formed after the replicator, which perhaps by then was RNA (the RNA world hypothesis), along with its replicating apparatus and other biomolecules. Initial protocells may have simply burst when they grew too large; the scattered contents may then have recolonized other “bubbles”. Proteins that stabilized the membrane, or that later assisted in an orderly division, would have promoted the proliferation of those cell lines.

RNA is a likely candidate for an early replicator, because it can both store genetic information and catalyze reactions. At some point DNA took over the genetic storage role from RNA, and proteins known as enzymes took over the catalysis role, leaving RNA to transfer information, synthesize proteins and regulate the process. There is increasing belief that these early cells evolved in association with undersea volcanic vents known as *black smokers* or even hot, deep rocks.

It is believed that of this multiplicity of protocells, only one line survived. Current phylogentic evidence suggests that the last universal common ancestor (LUCA)lived during the early Archean eon, perhaps roughly 3.5 Ga or earlier. This LUCA cell is the ancestor of all life on Earth today. It was probably a prokaryote, possessing a cell membrane and probably ribosomes, but lacking a nucleus or membrane-bound organelles such as mitochondria or chloroplasts.

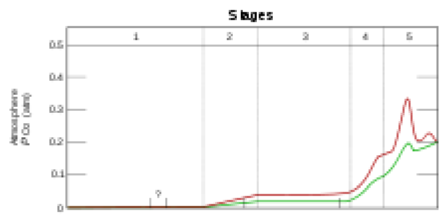
Like all modern cells, it used DNA as its genetic code, RNA for information transfer and protein synthesis, and enzymes to catalyze reactions. Some scientists believe that instead of a single organism being the last universal common ancestor, there were populations of organisms exchanging genes in lateral gene transfer.

Proterozoic eon

The Proterozoic is the eon of Earth's history that lasted from 2.5 Ga to 542 Ma. In this time span, the cratons grew into continents with modern sizes. For the first time plate tectonics took place in a modern sense. The change to an oxygen-rich atmosphere was a crucial development. Life developed from prokaryotes into eukaryotes and multicellular forms. The Proterozoic saw a couple of severe ice ages called snowball Earths. After the end of the last Snowball Earth about 600 Ma, the evolution of life on Earth accelerated. About 580 Ma, the Ediacara biota formed the prelude for the Cambrian Explosion.

3.0 Ga: Oxygen revolution

The harnessing of the sun's energy led to several major changes in life on Earth.



Graph showing range of estimated partial pressure of atmospheric oxygen through geologic time

A banded iron formation from the 3.15 Ga Moories Group, Barberton Greenstone Belt, South Africa. Red layers represent the times when oxygen was available, gray layers were formed in anoxic circumstances.

The first cells were likely heterotrophs, using surrounding organic molecules (including those from other cells) as raw material and an energy source. As the food supply diminished, a new strategy evolved in some cells. Instead of relying on the diminishing amounts of free-existing organic molecules, these cells adopted sunlight as an energy source. Estimates vary, but by about 3 Ga, something similar to modern oxygenic photosynthesis had probably developed, which made the sun's energy available not only to autotrophs but also to the heterotrophs that consumed them. This type of photosynthesis, which became by far the most common, used the abundant carbon dioxide and water as raw materials and, with the energy of sunlight, produced energy-rich organic molecules (carbohydrates).

Moreover, oxygen was released as a waste product of the photosynthesis. At first, it became bound up with limestone, iron, and other minerals. There is substantial proof of this in iron-oxide rich layers in geological strata that correspond with this period. The reaction of the minerals with oxygen would have turned the oceans green. When most of the exposed readily reacting minerals were oxidized, oxygen finally began to accumulate in the atmosphere. Though each cell only produced a minute amount of oxygen, the combined metabolism of many cells over a vast time transformed Earth's atmosphere to its current state. Among the oldest examples of oxygen-producing lifeforms are fossil stromatolites. This was Earth's third atmosphere.

Some of the oxygen was stimulated by incoming ultraviolet radiation to form ozone, which collected in a layer near the upper part of the atmosphere. The ozone layer absorbed, and still absorbs, a significant amount of the ultraviolet radiation that once had passed through the atmosphere. It allowed cells to colonize the surface of the ocean and eventually the land: without the ozone layer, ultraviolet radiation bombarding land and sea would have caused unsustainable levels of mutation in exposed cells.

Photosynthesis had another, major, and world-changing impact. Oxygen was toxic; probably much life on Earth died out as its levels rose in what is known as the "oxygen catastrophe". Resistant forms survived and thrived, and some developed the ability to use oxygen to increase their metabolism and obtain more energy from the same food.

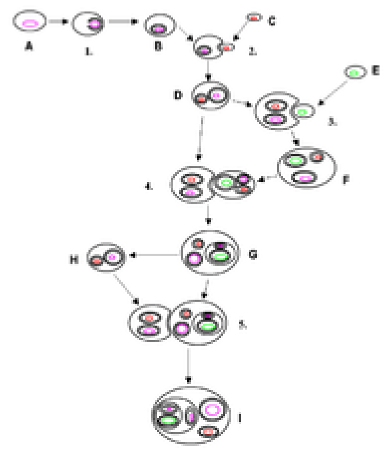
2.3 Ga: Snowball Earth and ozone layer

An oxygen-rich atmosphere had two principal advantages for life. Organisms not using oxygen for their metabolism, such as anaerobe bacteria, base their metabolism on fermentation. The abundance of oxygen makes respiration possible, a much more effective energy source for life than fermentation. The second advantage of an oxygen-rich atmosphere is that oxygen forms ozone in the higher atmosphere, causing the emergence of the Earth's ozone layer. The ozone layer protects the Earth's surface from ultraviolet radiation, which is harmful for life. Without the ozone layer, the development of more complex life later on would probably have been impossible.

The natural evolution of the Sun made it progressively more luminous during the Archaean and Proterozoic eons; the Sun's luminosity increases 6% every billion years. As a result, the Earth began to receive more heat from the Sun in the Proterozoic eon. However, the Earth did not get warmer. Instead, the geological record seems to suggest it cooled dramatically during the early Proterozoic. Glacial deposits found in all cratons show that about 2.3 Ga, the Earth underwent its first big ice age (the Makganyene ice age). Some scientists suggest this and following Proterozoic ice ages were so severe that the planet was totally frozen over from the poles to the equator, a hypothesis called Snowball Earth. Not all geologists agree with this scenario and older, Archaean ice ages have been postulated, but the ice age 2.3 Ga is the first such event for which the evidence is widely accepted.

The ice age around 2.3 Ga could have been directly caused by the increased oxygen concentration in the atmosphere, which caused the decrease of methane (CH_4) in the atmosphere. Methane is a strong greenhouse gas, but with oxygen it reacts to form CO_2, a less effective greenhouse gas. When free oxygen became available in the atmosphere, the concentration of methane could have decreased dramatically, enough to counter the effect of the increasing heat flow from the Sun.

2.0 Ga: Proterozoic development of life

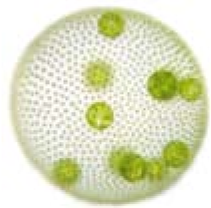


Some of the pathways by which the various endosymbionts might have arisen

Modern taxonomy classifies life into three domains. The time of the origin of these domains is uncertain. The Bacteria domain probably first split off from the other forms of life (sometimes called Neomura), but this supposition is controversial. Soon after this, by 2 Ga, the Neomura split into the Archaea and the Eukarya. Eukaryotic cells (Eukarya) are larger and more complex than prokaryotic cells (Bacteria and Archaea), and the origin of that complexity is only now becoming known.

Around this time, the first proto-mitochondrion was formed. A bacterial cell related to today's *Rickettsia*, which had evolved to metabolize oxygen, entered a larger prokaryotic cell, which lacked that capability. Perhaps the large cell attempted to ingest the smaller one but failed (possibly due to the evolution of prey defenses). The smaller cell may have tried to parasitize the larger one. In any case, the smaller cell survived inside the larger cell. Using oxygen, it metabolized the larger cell's waste products and derived more energy. Some of this excess energy was returned to the host. The smaller cell replicated inside the larger one. Soon, a stable symbiosis developed between the large cell and the smaller cells inside it. Over time, the host cell acquired some of the genes of the smaller cells, and the two kinds became dependent on each other: the larger cell could not survive without the energy produced by the smaller ones, and these in turn could not survive without the raw materials provided by the larger cell. The whole cell is now considered a single organism, and the smaller cells are classified as organelles called mitochondria.

A similar event occurred with photosynthetic cyanobacteria entering large heterotrophic cells and becoming chloroplasts. Probably as a result of these changes, a line of cells capable of photosynthesis split off from the other eukaryotes more than 1 billion years ago. There were probably several such inclusion events, as the figure at right suggests. Besides the well-established endosymbiotic theory of the cellular origin of mitochondria and chloroplasts, it has been suggested that cells led to peroxisomes, spirochetes led to cilia and flagella, and that perhaps a DNA virus led to the cell nucleus, though none of these theories is widely accepted.

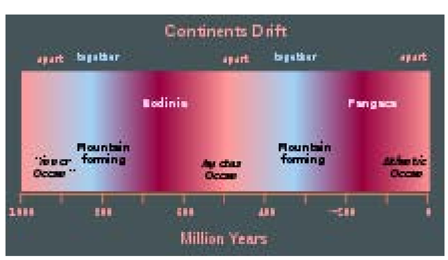


Green algae of the genus *Volvox* are believed to be similar to the first multicellular plants

Archaeans, bacteria, and eukaryotes continued to diversify and to become more complex and better adapted to their environments. Each domain repeatedly split into multiple lineages, although little is known about the history of the archaea and bacteria. Around 1.1 Ga, the supercontinent Rodinia was assembling. The plant, animal, and fungi lines had all split, though they still existed as solitary cells. Some of these lived in colonies, and gradually some division of labor began to take place; for instance, cells on the periphery might have started to assume different roles from those in the interior. Although the division between a colony with specialized cells and a multicellular organism is not always clear, around 1 billion years ago the first multicellular plants emerged, probably green algae. Possibly by around 900 Ma true multicellularity had also evolved in animals.

At first it probably resembled today's sponges, which have totipotent cells that allow a disrupted organism to reassemble itself. As the division of labor was completed in all lines of multicellular organisms, cells became more specialized and more dependent on each other; isolated cells would die.

Supercontinents in the Proterozoic era



Wilson cycle timeline from 1 Ga, depicting Rodinia and Pangaea supercontinent formation and separation

When the theory of Plate tectonics was developed around 1960, geologists began to reconstruct the movements and positions of the continents in the past. This appeared relatively easy until about 250 million years ago, when all continents were united in what is called the "supercontinent" Pangaea. Before that time, reconstructions cannot rely on apparent similarities in coastlines or ages of oceanic crust, but only on geologic observations and paleomagnetic data.

Throughout the history of the Earth, there have been times when the continental mass came together to form a supercontinent, followed by the break-up of the supercontinent and new continents moving apart again. This repetition of tectonic events is called a Wilson cycle. The further back in time, the scarcer and harder to interpret the data get. It is at least clear that, about 1000 to 830 Ma, most continental mass was united in the supercontinent Rodinia. Rodinia was not the first supercontinent; it formed at ~1.0 Ga by accretion and collision of fragments produced by breakup of the older supercontinent, called Nuna or Columbia, which was assembled by global-scale 2.0-1.8 Ga collisional events. This means plate tectonic processes similar to today's must have been active during the Proterozoic.

After the break-up of Rodinia about 800 Ma, it is possible the continents joined again around 550 Ma. The hypothetical supercontinent is sometimes referred to as Pannotia or Vendia. The evidence for it is a phase of continental collision known as the Pan-African orogeny, which joined the continental masses of current-day Africa, South-America, Antarctica and Australia. It is extremely likely, however, that the aggregation of continental masses was not completed, since a continent called Laurentia (roughly equivalent to current-day North America) had already started breaking off around 610 Ma. It is at least certain that by the end of the Proterozoic eon, most of the continental mass lay united in a position around the south pole.

Late Proterozoic climate and life

A 580 million year old fossil of *Spriggina floundensi*, an animal from the Ediacaran period. Such life forms could have been ancestors to the many new forms that origined in the Cambrian Explosion.

The end of the Proterozoic saw at least two Snowball Earths, so severe that the surface of the oceans may have been completely frozen. This happened about 710 and 640 Ma, in the Cryogenian period. These severe glaciations are less easy to explain than the early Proterozoic Snowball Earth. Most paleoclimatologists think the cold episodes had something to do with the formation of the supercontinent Rodinia. Because Rodinia was centered on the equator, rates of chemical weathering increased and carbon dioxide (CO_2) was taken from the atmosphere. Because CO_2 is an important greenhouse gas, climates cooled globally.

In the same way, during the Snowball Earths most of the continental surface was in permafrost, which decreased chemical weathering again, leading to the end of the glaciations. An alternative hypothesis is that enough carbon dioxide escaped through volcanic outgassing that the resulting greenhouse effect raised global temperatures. Increased volcanic activity resulted from the break-up of Rodinia at about the same time.

The Cryogenian period was followed by the Ediacaran period, which was characterized by a rapid development of new multicellular lifeforms. Whether there is a connection between the end of the severe ice ages and the increase in diversity of life is not clear, but it does not seem coincidental. The new forms of life, called Ediacara biota, were larger and more diverse than ever. Most scientists think some of them may have been the precursors of the new life forms of the following Cambrian period. Though the taxonomy of most Ediacaran life forms is unclear, some are proposed to have been ancestors of groups of modern life. Important developments were the origin of muscular and neural cells. None of the Ediacaran fossils had hard body parts like skeletons. These first appear after the boundary between the Proterozoic and Phanerozoic eons or Ediacaran and Cambrian periods.

Phanerozoic eon

Paleozoic era

The Paleozoic era (meaning: *era of old life forms*) was the first era of the Phanerozoic eon, lasting from 542 to 251 Ma. During the Paleozoic, many modern groups of life came into existence. Life colonized the land, first plants, then animals. Life usually evolved slowly. At times, however, there are sudden radiations of new species or mass extinctions. These bursts of evolution were often caused by unexpected changes in the environment resulting from natural disasters such as volcanic activity, meteorite impacts or climate changes.

The continents formed at the break-up of Pannotia and Rodinia at the end of the Proterozoic would slowly move together again during the Paleozoic. This would eventually result in phases of mountain building that created the supercontinent Pangaea in the late Paleozoic.

542 Ma: Cambrian explosion

Apparently, the rate of the evolution of life accelerated in the Cambrian period (542-488 Ma). The sudden emergence of many new species, phyla, and forms in this period is called the Cambrian Explosion. The biological formenting in the Cambrian Explosion was unpreceded before and since that time. Whereas the Ediacaran life forms appear yet primitive and not easy to put in any modern group, at the end of the Cambrian most modern phyla were already present. The development of hard body parts such as shells, skeletons or exoskeletons in animals like molluscs, echinoderms, crinoids and arthropods (a well-known group of arthropods from the lower Paleozoic are the trilobites) made the preservation and fossilisation of such life forms easier than those of their Proterozoic ancestors. For this reason, much more is known about life in and after the Cambrian than about that of older periods. The boundary between the Cambrian and Ordovician (the following period, 488-444 Ma) is characterized by a large mass-extinction, in which some of the new groups disappeared altogether. Some of these Cambrian groups appear complex but are quite different from modern life; examples are *Anomalocaris* and *Haikouichthys*.

During the Cambrian, the first vertebrate animals, among them the first fishes, had appeared. A creature that could have been the ancestor of the fishes, or was probably closely related to it, was *Pikaia*. It had a primitive notochord, a structure that could have developed into a vertebral column later. The first fishes with jaws (Gnathostomata) appeared during the Ordovician. The colonisation of new niches resulted in massive body sizes. In this way, fishes with increasing sizes evolved during the early Paleozoic, such as the titanic placoderm *Dunkleosteus*, which could grow 7 meters long.

Paleozoic tectonics, paleogeography and climate

At the end of the Proterozoic, the supercontinent Pannotia had broken apart in the smaller continents Laurentia, Baltica, Siberia and Gondwana. During periods when continents move apart, more oceanic crust is formed by volcanic activity. Because young volcanic crust is relatively hotter and less dense than old oceanic crust, the ocean floors will rise during such periods. This causes the sea level to rise. Therefore, in the first half of the Paleozoic, large areas of the continents were below sea level.

Early Paleozoic climates were warmer than today, but the end of the Ordovician saw a short ice age during which glaciers covered the south pole, where the huge continent Gondwana was situated. Traces of glaciation from this period are only found on former Gondwana. During the Late Ordovician ice age, a number of mass extinctions took place, in which many brachiopods, trilobites, Bryozoa and corals disappeared. These marine species could probably not contend with the decreasing temperature of the sea water. After the extinctions new species evolved, more diverse and better adapted. They would fill the niches left by the extinct species.

The continents Laurentia and Baltica collided between 450 and 400 Ma, during the Caledonian Orogeny, to form Laurussia. Traces of the mountain belt which resulted from this collision can be found in Scandinavia, Scotland and the northern Appalachians. In the Devonian period (416-359 Ma) Gondwana and Siberia began to move towards Laurussia. The collision of Siberia with Laurussia caused the Uralian Orogeny, the collision of Gondwana with Laurussia is called the Variscan or Hercynian Orogeny in Europe or the Alleghenian Orogeny in North America. The latter phase took place during the Carboniferous period (359-299 Ma) and resulted in the formation of the last supercontinent, Pangaea.

Colonization of land

For most of Earth's history, there were no multicellular organisms on land. Parts of the surface may have vaguely resembled this view of Mars.

Oxygen accumulation from photosynthesis resulted in the formation of an ozone layer that absorbed much of Sun's ultraviolet radiation, meaning unicellular organisms that reached land were less likely to die, and prokaryotes began to multiply and become better adapted to survival out of the water. A variety of prokaryote lineages had probably colonized the land as early as 2.6 Ga even before the origin of the eukaryotes. For a long time, the land remained barren of multicellular organisms. The supercontinent Pannotia formed around 600 Ma and then broke apart a short 50 million years later. Fish, the earliest vertebrates, evolved in the oceans around 530 Ma. A major extinction event occurred near the end of the Cambrian period, which ended 488 Ma.

Several hundred million years ago, plants (probably resembling algae) and fungi started growing at the edges of the water, and then out of it. The oldest fossils of land fungi and plants date to 480–460 Ma, though molecular evidence suggests the fungi may have colonized the land as early as 1000 Ma and the plants 700 Ma. Initially remaining close to the water's edge, mutations and variations resulted in further colonization of this new environment. The timing of the first animals to leave the oceans is not precisely known: the oldest clear evidence is of arthropods on land around 450 Ma, perhaps thriving and becoming better adapted due to the vast food source provided by the terrestrial plants. There is also some unconfirmed evidence that arthropods may have appeared on land as early as 530 Ma.

At the end of the Ordovician period, 440 Ma, additional extinction events occurred, perhaps due to a concurrent ice age. Around 380 to 375 Ma, the first tetrapods evolved from fish. It is thought that perhaps fins evolved to become limbs which allowed the first tetrapods to lift their heads out of the water to breathe air. This would allow them to live in oxygen-poor water or pursue small prey in shallow water. They may have later ventured on land for brief periods. Eventually, some of them became so well adapted to terrestrial life that they spent their adult lives on land, although they hatched in the water and returned to lay their eggs. This was the origin of the amphibians. About 365 Ma, another period of extinction occurred, perhaps as a result of global cooling. Plants evolved seeds, which dramatically accelerated their spread on land, around this time (by approximately 360 Ma).

Pangaea, the most recent supercontinent, existed from 300 to 180 Ma. The outlines of the modern continents and other land masses are indicated on this map.

Some 20 million years later (340 Ma), the amniotic egg evolved, which could be laid on land, giving a survival advantage to tetrapod embryos. This resulted in the divergence of amniotes from amphibians. Another 30 million years (310 Ma) saw the divergence of the synapsids (including mammals) from the sauropsids (including birds and reptiles). Other groups of organisms continued to evolve, and lines diverged—in fish, insects, bacteria, and so on—but less is known of the details. The most recent hypothesized supercontinent, called Pangaea, formed 300 Ma.

Mesozoic era

The most severe extinction event to date took place 250 Ma, at the boundary of the Permian and Triassic periods; 95% of life on Earth died out and started the Mesozoic era (meaning middle life) that spanned 187 million years. This extinction event was possibly caused by the Siberian Traps volcanic event, an asteroid impact, methane hydrate gasification, sea level fluctuations, a major anoxic event, other events, or some combination of these events. Either the proposed Wilkes Land crater in Antarctica or Bedout structure off the northwest coast of Australia may indicate an impact connection with the Permian-Triassic extinction. But it remains uncertain whether either these or other proposed Permian-Triassic boundary craters are either real impact craters or even contemporaneous with the Permian-Triassic extinction event. Life persevered, and around 230 Ma, dinosaurs split off from their reptilian ancestors. An extinction event between the Triassic and Jurassic periods 200 Ma spared many of the dinosaurs, and they soon became dominant among the vertebrates. Though some of the mammalian lines began to separate during this period, existing mammals were probably all small animals resembling shrews.

By 180 Ma, Pangaea broke up into Laurasia and Gondwana. The boundary between avian and non-avian dinosaurs is not clear, but *Archaeopteryx*, traditionally considered one of the first birds, lived around 150 Ma. The earliest evidence for the angiosperms evolving flowers is during the Cretaceous period, some 20 million years later (132 Ma). Competition with birds drove many pterosaurs to extinction and the dinosaurs were probably already in decline when, 65 Ma, a 10-kilometre (6.2 mi) meteorite probably struck Earth just off the Yucatán Peninsula where the Chicxulub crater is today. This ejected vast quantities of particulate matter and vapor into the air that occluded sunlight, inhibiting photosynthesis. Most large animals, including the non-avian dinosaurs, became extinct, marking the end of the Cretaceous period and Mesozoic era. Thereafter, in the Paleocene epoch, mammals rapidly diversified, grew larger, and became the dominant vertebrates. Perhaps a couple of million years later (around 63 Ma), the last common ancestor of primates lived. By the late Eocene epoch, 34 Ma, some terrestrial mammals had returned to the oceans to become animals such as *Basilosaurus* which eventually led to dolphins and baleen whales.

Cenozoic era (Recent life)

2 Ma: Human evolution

Australopithecus africanus, an early hominid

A small African ape living around 6 Ma was the last animal whose descendants would include both modern humans and their closest relatives, the bonobo and chimpanzees. Only two branches of its family tree have surviving descendants. Very soon after the split, for reasons that are still debated, apes in one branch developed the ability to walk upright. Brain size increased rapidly, and by 2 Ma, the first animals classified in the genus *Homo* had appeared. Of course, the line between different species or even genera is somewhat arbitrary as organisms continuously change over generations. Around the same time, the other branch split into the ancestors of the common chimpanzee and the ancestors of the bonobo as evolution continued simultaneously in all life forms.

The ability to control fire probably began in *Homo erectus* (or *Homo ergaster*), probably at least 790,000 years ago but perhaps as early as 1.5 Ma. In addition, it has sometimes suggested that the use and discovery of controlled fire may even predate *Homo erectus*. Fire was possibly used by the early Lower Paleolithic (Oldowan) hominid *Homo habilis* or strong australopithecines such as *Paranthropus.*

It is more difficult to establish the origin of language; it is unclear whether *Homo erectus* could speak or if that capability had not begun until *Homo sapiens*. As brain size increased, babies were born earlier, before their heads grew too large to pass through the pelvis. As a result, they exhibited more plasticity, and thus possessed an increased capacity to learn and required a longer period of dependence. Social skills became more complex, language became more sophisticated, and tools became more elaborate. This contributed to further cooperation and intellectual development. Modern humans (*Homo sapiens*) are believed to have originated somewhere around 200,000 years ago or earlier in Africa; the oldest fossils date back to around 160,000 years ago.

The first humans to show signs of spirituality are the Neanderthals (usually classified as a separate species with no surviving descendants); they buried their dead, often apparently with food or tools. However, evidence of more sophisticated beliefs, such as the early Cro-Magnon cave paintings (probably with magical or religious significance) did not appear until some 32,000 years ago. Cro-Magnons also left behind stone figurines such as Venus of Willendorf, probably also signifying religious belief. By 11,000 years ago, *Homo sapiens* had reached the southern tip of South America, the last of the uninhabited continents (except for Antarctica, which remained undiscovered until 1820 AD). Tool use and communication continued to improve, and interpersonal relationships became more intricate.

10,000 ya: Civilization

Vitruvian Man by Leonardo da Vinci epitomizes the advances in art and science seen during the Renaissance.

Throughout more than 90% of its history, *Homo sapiens* lived in small bands as nomadic hunter-gatherers. As language became more complex, the ability to remember and communicate information resulted in a new replicator: the meme. Ideas could be exchanged quickly and passed down the generations.

Cultural evolution quickly outpaced biological evolution, and history proper began. Somewhere between 8500 and 7000 BC, humans in the Fertile Crescent in Middle East began the systematic husbandry of plants and animals: agriculture. This spread to neighboring regions, and developed independently elsewhere, until most *Homo sapiens* lived sedentary lives in permanent settlements as farmers.

Not all societies abandoned nomadism, especially those in isolated areas of the globe poor in domesticable plant species, such as Australia. However, among those civilizations that did adopt agriculture, the relative stability and increased productivity provided by farming allowed the population to expand.

Agriculture had a major impact; humans began to affect the environment as never before. Surplus food allowed a priestly or governing class to arise, followed by increasing division of labor. This led to Earth's first civilization at Sumer in the Middle East, between 4000 and 3000 BC. Additional civilizations quickly arose in ancient Egypt, at the Indus River valley and in China.

Starting around 3000 BC, Hinduism, one of the oldest religions still practiced today, began to take form. Others soon followed. The invention of writing enabled complex societies to arise: record-keeping and libraries served as a storehouse of knowledge and increased the cultural transmission of information. Humans no longer had to spend all their time working for survival—curiosity and education drove the pursuit of knowledge and wisdom.

Various disciplines, including science (in a primitive form), arose. New civilizations sprang up, traded with one another, and fought for territory and resources. Empires soon began to develop. By around 500 BC, there were empires in the Middle East, Iran, India, China, and Greece, on nearly equal footing; at times one empire expanded, only to decline or be driven back later.

In the fourteenth century, the Renaissance began in Italy with advances in religion, art, and science. European civilization began to change beginning in 1500, leading to the scientific and industrial revolutions. That continent began to exert political and cultural dominance over human societies around the planet. From 1914 to 1918 and 1939 to 1945, nations around the world were embroiled in world wars.

Established following World War I, the League of Nations was a first step in establishing international institutions to settle disputes peacefully. After failing to prevent World War II, it was replaced by the United Nations. In 1992, several European nations joined in the European Union. As transportation and communication improved, the economies and political affairs of nations around the world have become increasingly intertwined. This globalization has often produced both conflict and collaboration.

75 ya: Recent events

Four and a half billion years after the planet's formation, Earth's life broke free of the biosphere. For the first time in history, Earth was viewed from space.

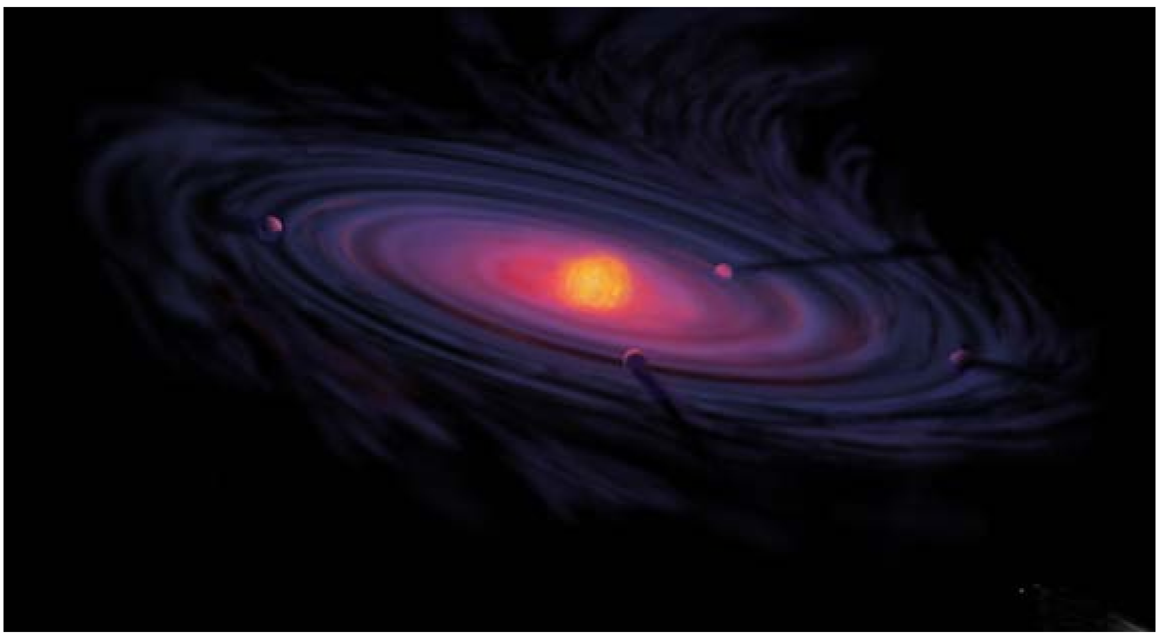
Change has continued at a rapid pace from the mid-1940s to today. Technological developments include nuclear weapons, computers, genetic engineering, and nanotechnology. Economic globalization spurred by advances in communication and transportation technology has influenced everyday life in many parts of the world.

Cultural and institutional forms such as democracy, capitalism, and environmentalism have increased influence. Major concerns and problems such as disease, war, poverty, violent radicalism, and recently, human-caused climate change have risen as the world population increases.

In 1957, the Soviet Union launched the first artificial satellite into orbit and, soon afterward, Yuri Gagarin became the first human in space. Neil Armstrong, an American, was the first to set foot on another astronomical object, the Moon. Unmanned probes have been sent to all the known planets in the solar system, with some (such as Voyager) having left the solar system. The Soviet Union and the United States were the earliest leaders in space exploration in the 20th Century. Five space agencies, representing over fifteen countries, have worked together to build the International Space Station. Aboard it, there has been a continuous human presence in space since 2000.

Chapter- 2

Formation and Evolution of the Solar System



Artist's conception of a protoplanetary disk

The formation and evolution of the Solar System is estimated to have begun 4.568 billion years ago with the gravitational collapse of a small part of a giant molecular cloud.

Most of the collapsing mass collected in the centre, forming the Sun, while the rest flattened into a protoplanetary disk out of which the planets, moons, asteroids, and other small Solar System bodies formed.

This widely accepted model, known as the nebular hypothesis, was first developed in the 18th century by Emanuel Swedenborg, Immanuel Kant, and Pierre-Simon Laplace. Its subsequent development has interwoven a variety of scientific disciplines including astronomy, physics, geology, and planetary science. Since the dawn of the space age in the 1950s and the discovery of extrasolar planets in the 1990s, the models have been both challenged and refined to account for new observations.

The Solar System has evolved considerably since its initial formation. Many moons have formed from circling discs of gas and dust around their parent planets, while other moons are believed to have formed independently and later been captured by their planets. Still others, as the Earth's Moon, may be the result of giant collisions. Collisions between bodies have occurred continually up to the present day and have been central to the evolution of the solar system. The positions of the planets often shifted, and planets have switched places. This planetary migration now is believed to have been responsible for much of the Solar System's early evolution.

In roughly 5 billion years, the Sun will cool and expand outward to many times its current diameter (becoming a red giant), before casting off its outer layers as a planetary nebula, and leaving behind a stellar corpse known as a white dwarf. In the far distant future, the gravity of passing stars gradually will whittle away at the Sun's retinue of planets. Some planets will be destroyed, others ejected into interstellar space. Ultimately, over the course of trillions of years, it is likely that the Sun will be left alone with no bodies in orbit around it.

History

Pierre-Simon Laplace, one of the originators of the nebular hypothesis

Ideas concerning the origin and fate of the world date from the earliest known writings; however, for almost all of that time, there was no attempt to link such theories to the existence of a "Solar System", simply because it was not generally believed that the Solar System, in the sense we now understand it, existed. The first step toward a theory of Solar System formation and evolution was the general acceptance of heliocentrism, which placed the Sun at the centre of the system and the

Earth in orbit around it. This conception had gestated for millennia (philosophers such as Aristarchus of Samos had suggested it as early as 600 BC), but was widely accepted only by the end of the 17th century. The first recorded use of the term "Solar System" dates from 1704.

The current standard theory for Solar System formation, the nebular hypothesis, has fallen into and out of favour since its formulation by Emanuel Swedenborg, Immanuel Kant, and Pierre-Simon Laplace in the 18th century. The most significant criticism of the hypothesis was its apparent inability to explain the Sun's relative lack of angular momentum when compared to the planets. However, since the early 1980s studies of young stars have shown them to be surrounded by cool discs of dust and gas, exactly as the nebular hypothesis predicts, which has led to its re-acceptance.

Understanding of how the Sun will continue to evolve required an understanding of the source of its power. Arthur Stanley Eddington's confirmation of Albert Einstein's theory of relativity led to his realisation that the Sun's energy comes from nuclear fusion reactions in its core. In 1935, Eddington went further and suggested that other elements also might form within stars. Fred Hoyle elaborated on this premise by arguing that evolved stars called red giants created many elements heavier than hydrogen and helium in their cores. When a red giant finally casts off its outer layers, these elements would then be recycled to form other star systems.

Formation

Pre-solar nebula

The nebular hypothesis maintains that the Solar System formed from the gravitational collapse of a fragment of a giant molecular cloud. The cloud itself had a size of about 20 pc, while the fragments were roughly 1 pc (several light-years) across. The further collapse of the fragments led to the formation of dense cores 0.01–0.1 pc (2,000–20,000 AU) in size. One of these collapsing fragments (known as the *pre-solar nebula*) would form what became the Solar System. The composition of this region with a mass just over that of the Sun was about the same as that of the Sun today, with hydrogen, along with helium and trace amounts of lithium produced by Big Bang nucleosynthesis, forming about 98% of its mass. The remaining 2% of the mass consisted of heavier elements that were created by nucleosynthesis in earlier generations of stars. Late in the life of these stars, they ejected heavier elements into the interstellar medium.

Hubble image of protoplanetary discs in the Orion nebula, a light-years-wide "stellar nursery" likely very similar to the primordial nebula from which our Sun formed

Studies of ancient meteorites reveal traces of stable daughter nuclei of short-lived isotopes, such as iron-60, that only form in exploding, short-lived stars. This indicates that one or more supernovae occurred near the Sun while it was forming. A shock wave from a supernova may have triggered the formation of the Sun by creating regions of over-density within the cloud, causing these regions to collapse. Because only massive, short-lived stars produce supernovae, the Sun must have formed in a large star-forming region that produced massive stars, possibly similar to the Orion nebula. Studies of the structure of the Kuiper belt and of anomalous materials within it suggest that the Sun formed within a cluster of stars with a diameter of between 6.5 and 19.5 light-years and a collective mass equivalent to 3,000 Suns. Several simulations of our young Sun interacting with close-passing stars over the first 100 million years of its life produce anomalous orbits observed in the outer Solar System, such as detached objects.

Because of the conservation of angular momentum, the nebula spun faster as it collapsed. As the material within the nebula condensed, the atoms within it began to collide with increasing frequency, converting their kinetic energy into heat. The centre, where most of the mass collected, became increasingly hotter than the surrounding disc. Over about 100,000 years, the competing forces of gravity, gas pressure, magnetic fields, and rotation caused the contracting nebula to flatten into a spinning protoplanetary disc with a diameter of ~200 AU and form a hot, dense protostar (a star in which hydrogen fusion has not yet begun) at the centre.

At this point in its evolution, the Sun is believed to have been a T Tauri star. Studies of T Tauri stars show that they are often accompanied by discs of pre-planetary matter with masses of 0.001–0.1 solar masses. These discs extend to several hundred AU—the Hubble Space Telescope has observed protoplanetary discs of up to 1000 AU in diameter in star-forming regions such as the Orion Nebula—and are rather cool, reaching only a thousand kelvins at their hottest. Within 50 million years, the temperature and pressure at the core of the Sun became so great that its hydrogen began to fuse, creating an internal source of energy that countered gravitational contraction until hydrostatic equilibrium was achieved. This marked the Sun's entry into the prime phase of its life, known as the main sequence. Main sequence stars derive energy from the fusion of hydrogen into helium in their cores. The Sun remains a main sequence star today.

Formation of planets

Artist's conception of the solar nebula

The various planets are thought to have formed from the *solar nebula*, the disc-shaped cloud of gas and dust left over from the Sun's formation. The currently accepted method by which the planets formed is known as accretion, in which the planets began as dust grains in orbit around the central protostar. Through direct contact, these grains formed into clumps up to 200 metres in diameter, which in turn collided to form larger bodies (planetesimals) of ~10 kilometres (km) in size. These gradually increased through further collisions, growing at the rate of centimetres per year over the course of the next few million years.

The inner Solar System, the region of the Solar System inside 4 AU, was too warm for volatile molecules like water and methane to condense, so the planetesimals that formed there could only form from compounds with high melting points, such as metals (like iron, nickel, and aluminium) and rocky silicates. These rocky bodies would become the terrestrial planets (Mercury, Venus, Earth, and Mars). These compounds are quite rare in the universe, comprising only 0.6% of the mass of the nebula, so the terrestrial planets could not grow very large. The terrestrial embryos grew to about 0.05 Earth masses and ceased accumulating matter about 100,000 years after the formation of the Sun; subsequent collisions and mergers between these planet-sized bodies allowed terrestrial planets to grow to their present sizes.

When the terrestrial planets were forming, they remained immersed in a disk of gas and dust. The gas was partially supported by pressure and so did not orbit the Sun as rapidly as the planets. The resulting drag caused a transfer of angular momentum, and as a result the planets gradually migrated to new orbits. Models show that temperature variations in the disk governed this rate of migration, but the net trend was for the inner planets to migrate inward as the disk dissipated, leaving the planets in their current orbits.

The gas giant planets (Jupiter, Saturn, Uranus, and Neptune) formed further out, beyond the frost line, the point between the orbits of Mars and Jupiter where the material is cool enough for volatile icy compounds to remain solid. The ices that formed the Jovian planets were more abundant than the metals and silicates that formed the terrestrial planets, allowing the Jovian planets to grow massive enough to capture hydrogen and helium, the lightest and most abundant elements. Planetesimals beyond the frost line accumulated up to four Earth masses within about 3 million years. Today, the four gas giants comprise just under 99% of all the mass orbiting the Sun. Theorists believe it is no accident that Jupiter lies just beyond the frost line. Because the frost line accumulated large amounts of water via evaporation from infalling icy material, it created a region of lower pressure that increased the speed of orbiting dust particles and halted their motion toward the Sun. In effect, the frost line acted as a barrier that caused material to accumulate rapidly at ~5 AU from the Sun. This excess material coalesced into a large embryo of about 10 Earth masses, which then began to grow rapidly by swallowing hydrogen from the surrounding disc, reaching 150 Earth masses in only another 1000 years and finally topping out at 318 Earth masses. Saturn may owe its substantially lower mass simply to having formed a few million years after Jupiter, when there was less gas available to consume.

T Tauri stars like the young Sun have far stronger stellar winds than more stable, older stars. Uranus and Neptune are believed to have formed after Jupiter and Saturn did, when the strong solar wind had blown away much of the disc material. As a result, the planets accumulated little hydrogen and helium—not more than 1 Earth mass each. Uranus and Neptune are sometimes referred to as failed cores. The main problem with formation theories for these planets is the timescale of their formation. At the current locations it would have taken a hundred million years for their cores to accrete. This means that Uranus and Neptune probably formed closer to the Sun—near or even between Jupiter and Saturn—and later migrated outward. Motion in the planetesimal era was not all inward toward the Sun; the *Stardust* sample return from Comet Wild 2 has suggested that materials from the early formation of the Solar System migrated from the warmer inner Solar System to the region of the Kuiper belt.

After between three and ten million years, the young Sun's solar wind would have cleared away all the gas and dust in the protoplanetary disc, blowing it into interstellar space, thus ending the growth of the planets.

Subsequent evolution

The giant impact believed to have formed the Moon

The planets were originally believed to have formed in or near their current orbits. However, this view underwent radical change during the late 20th and early 21st centuries. Currently, it is believed that the Solar System looked very different after its initial formation: several objects at least as massive as Mercury were present in the inner Solar System, the outer Solar System was much more compact than it is now, and the Kuiper belt was much closer to the Sun.

Terrestrial planets

At the end of the planetary formation epoch the inner Solar System was populated by 50–100 Moon- to Mars-sized planetary embryos. Further growth was possible only because these bodies collided and merged, which took less than 100 million years. These objects would have gravitationally interacted with one another, tugging at each other's orbits until they collided, growing larger until the four terrestrial planets we know today took shape. One such giant collision is believed to have formed the Moon, while another removed the outer envelope of the young Mercury.

One unresolved issue with this model is that it cannot explain how the initial orbits of the proto-terrestrial planets, which would have needed to be highly eccentric to collide, produced the remarkably stable and near-circular orbits the terrestrial planets possess today. One hypothesis for this "eccentricity dumping" is that the terrestrials formed in a disc of gas still not expelled by the Sun. The "gravitational drag" of this residual gas would have eventually lowered the planets' energy, smoothing out their orbits. However, such gas, if it existed, would have prevented the terrestrials' orbits from becoming so eccentric in the first place. Another hypothesis is that gravitational drag occurred not between the planets and residual gas but between the planets and the remaining small bodies. As the large bodies moved through the crowd of smaller objects, the smaller objects, attracted by the larger planets' gravity, formed a region of higher density, a "gravitational wake", in the larger objects' path. As they did so, the increased gravity of the wake slowed the larger objects down into more regular orbits.

Asteroid belt

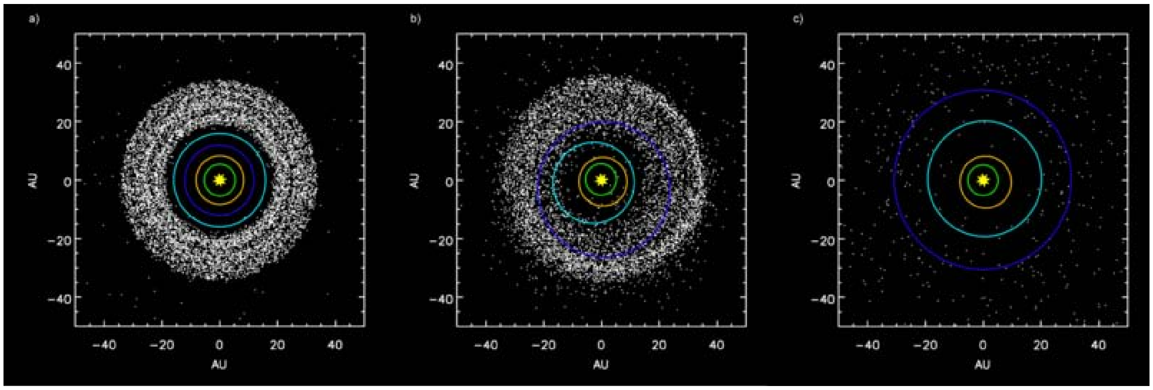
The outer edge of the terrestrial region, between 2 and 4 AU from Sun, is called the asteroid belt. The asteroid belt initially contained more than enough matter to form 2–3 Earth-like planets, and, indeed, a large number of planetesimals formed there. As with the terrestrials, planetesimals in this region later coalesced and formed 20–30 Moon- to Mars-sized planetary embryos; however, the proximity of Jupiter meant that after this planet formed, 3 million years after the Sun, the region's history changed dramatically. Orbital resonances with Jupiter and Saturn are particularly strong in the asteroid belt, and gravitational interactions with more massive embryos scattered many planetesimals into those resonances. Jupiter's gravity increased the velocity of objects within these resonances, causing them to shatter upon collision with other bodies, rather than accrete.

As Jupiter migrated inward following its formation, resonances would have swept across the asteroid belt, dynamically exciting the region's population and increasing their velocities relative to each other. The cumulative action of the resonances and the embryos either scattered the planetesimals away from the asteroid belt or excited their orbital inclinations and eccentricities. Some of those massive embryos too were ejected by Jupiter, while others may have migrated to the inner Solar System and played a role in the final accretion of the terrestrial planets. During this primary depletion period, the effects of the giant planets and planetary embryos left the asteroid belt with a total mass equivalent to less than 1% that of the Earth, composed mainly of small planetesimals. This is still 10–20 times more than the current mass in the main belt, which is about 1/2,000 the Earth's mass. A secondary depletion period that brought the asteroid belt down close to its present mass is believed to have followed when Jupiter and Saturn entered a temporary 2:1 orbital resonance (see below).

The inner Solar System's period of giant impacts probably played a role in the Earth acquiring its current water content ($\sim 6 \times 10^{21}$ kg) from the early asteroid belt. Water is too volatile to have been present at Earth's formation and must have been subsequently delivered from outer, colder parts of the Solar System. The water was probably delivered by planetary embryos and small planetesimals thrown out of the asteroid belt by Jupiter. A population of main-belt comets discovered in 2006 has been also suggested as a possible source for Earth's water. In contrast, comets from the Kuiper belt or farther regions delivered not more than about 6% of Earth's water. The panspermia hypothesis holds that life itself may have been deposited on Earth in this way, although this idea is not widely accepted.

Planetary migration

According to the nebular hypothesis, the outer two planets are in the "wrong place". Uranus and Neptune (known as the "ice giants") exist in a region where the reduced density of the solar nebula and longer orbital times render their formation highly implausible. The two are instead believed to have formed in orbits near Jupiter and Saturn, where more material was available, but to have migrated outward to their current positions over hundreds of millions of years.



Simulation showing outer planets and Kuiper belt: a) Before Jupiter/Saturn 2:1 resonance b) Scattering of Kuiper belt objects into the Solar System after the orbital shift of Neptune c) After ejection of Kuiper belt bodies by Jupiter

The migration of the outer planets is also necessary to account for the existence and properties of the Solar System's outermost regions. Beyond Neptune, the Solar System continues into the Kuiper belt, the scattered disc, and the Oort cloud, three sparse populations of small icy bodies thought to be the points of origin for most observed comets. At their distance from the Sun, accretion was too slow to allow planets to form before the solar nebula dispersed, and thus the initial disc lacked enough mass density to consolidate into a planet. The Kuiper belt lies between 30 and 55 AU from the Sun, while the farther scattered disc extends to over 100 AU, and the distant Oort cloud begins at about 50,000 AU. Originally, however, the Kuiper belt was much denser and closer to the Sun, with an outer edge at approximately 30 AU. Its inner edge would have been just beyond the orbits of Uranus and Neptune, which were in turn far closer to the Sun when they formed (most likely in the range of 15–20 AU), and in opposite locations, with Uranus farther from the Sun than Neptune.

After the formation of the Solar System, the orbits of all the giant planets continued to change slowly, influenced by their interaction with large number of remaining planetesimals. After 500–600 million years (about 4 billion years ago) Jupiter and Saturn fell into a 2:1 resonance; Saturn orbited the Sun once for every two Jupiter orbits. This resonance created a gravitational push against the outer planets, causing Neptune to surge past Uranus and plough into the ancient Kuiper belt. The planets scattered the majority of the small icy bodies inwards, while themselves moving outwards. These planetesimals then scattered off the next planet they encountered in a similar manner, moving the planets' orbits outwards while they moved inwards. This process continued until the planetesimals interacted with Jupiter, whose immense gravity sent them into highly elliptical orbits or even ejected them outright from the Solar System. This caused Jupiter to move slightly inward. Those objects scattered by Jupiter into highly elliptical orbits formed the Oort cloud; those objects scattered to a lesser degree by the migrating Neptune formed the current Kuiper belt and scattered disc. This scenario explains the Kuiper belt's and scattered disc's present low mass. Some of the scattered objects, including Pluto, became gravitationally tied to Neptune's orbit, forcing them into mean-motion resonances. Eventually, friction within the planetesimal disc made the orbits of Uranus and Neptune circular again.

In contrast to the outer planets, the inner planets are not believed to have migrated significantly over the age of the Solar System, because their orbits have remained stable following the period of giant impacts.

Late Heavy Bombardment and after

Meteor Crater in Arizona. Created 50,000 years ago by an impactor only 50m across, it is a stark reminder that the accretion of the Solar System is not over.

Gravitational disruption from the outer planets' migration would have sent large numbers of asteroids into the inner Solar System, severely depleting the original belt until it reached today's extremely low mass. This event may have triggered the Late Heavy Bombardment that occurred approximately 4 billion years ago, 500–600 million years after the formation of the Solar System. This period of heavy bombardment lasted several hundred million years and is evident in the cratering still visible on geologically dead bodies of the inner Solar System such as the Moon and Mercury. The oldest known evidence for life on Earth dates to 3.8 billion years ago—almost immediately after the end of the Late Heavy Bombardment.

Impacts are believed to be a regular (if currently infrequent) part of the evolution of the Solar System. That they continue to happen is evidenced by the collision of Comet Shoemaker-Levy 9 with Jupiter in 1994, the 2009 Jupiter impact event, and the impact feature Meteor Crater in Arizona. The process of accretion, therefore, is not complete, and may still pose a threat to life on Earth.

Over the course of the Solar System's evolution, comets were ejected out of the inner Solar System by the gravity of the giant planets, and sent thousands of AU outward to form the Oort cloud, a spherical outer swarm of cometary nuclei at the farthest extent of the Sun's gravitational pull. Eventually, after about 800 million years, the gravitational disruption caused by galactic tides, passing stars and giant molecular clouds began to deplete the cloud, sending comets into the inner Solar System. The evolution of the outer Solar System also appears to have been influenced by space weathering from the solar wind, micrometeorites, and the neutral components of the interstellar medium.

The evolution of the asteroid belt after Late Heavy Bombardment was mainly governed by collisions. Objects with large mass have enough gravity to retain any material ejected by a violent collision. In the asteroid belt this usually is not the case. As a result, many larger objects have been broken apart, and sometimes newer objects have been forged from the remnants in less violent collisions. Moons around some asteroids currently can only be explained as consolidations of material flung away from the parent object without enough energy to entirely escape its gravity.

Moons

Moons have come to exist around most planets and many other Solar System bodies. These natural satellites originated by one of three possible mechanisms:

- co-formation from a circum-planetary disc (only in the cases of the gas giants);
- formation from impact debris (given a large enough impact at a shallow angle); and
- capture of a passing object.

Jupiter and Saturn have a number of large moons, such as Io, Europa, Ganymede and Titan, which may have originated from discs around each giant planet in much the same way that the planets formed from the disc around the Sun. This origin is indicated by the large sizes of the moons and their proximity to the planet. These attributes are impossible to achieve via capture, while the gaseous nature of the primaries make formation from collision debris another impossibility. The outer moons of the gas giants tend to be small and have eccentric orbits with arbitrary inclinations. These are the characteristics expected of captured bodies. Most such moons orbit in the direction opposite the rotation of their primary. The largest irregular moon is Neptune's moon Triton, which is believed to be a captured Kuiper belt object.

Moons of solid Solar System bodies have been created by both collisions and capture. Mars's two small moons, Deimos and Phobos, are believed to be captured asteroids. The Earth's Moon is believed to have formed as a result of a single, large oblique collision. The impacting object likely had a mass comparable to that of Mars, and the impact probably occurred near the end of the period of giant impacts. The collision kicked into orbit some of the impactor's mantle, which then coalesced into the Moon. The impact was probably the last in the series of mergers that formed the Earth. It has been further hypothesized that the Mars-sized object may have formed at one of the stable Earth-Sun Lagrangian points (either L_4 or L_5) and drifted from its position. Pluto's moon Charon may also have formed by means of a large collision; the Pluto-Charon and Earth-Moon systems are the only two in the Solar System in which the satellite's mass is at least 1% that of the larger body.

Future

Astronomers estimate that the Solar System as we know it today will not change drastically until the Sun has fused all the hydrogen fuel in its core into helium, beginning its evolution from the main sequence of the Hertzsprung-Russell diagram and into its red giant phase. Even so, the Solar System will continue to evolve until then.

Long-term stability

The Solar System is chaotic, with the orbits of the planets open to long-term variations. One notable example of this chaos is the Neptune-Pluto system, which lies in a 3:2 orbital resonance.

Although the resonance itself will remain stable, it becomes impossible to predict the position of Pluto with any degree of accuracy more than 10–20 million years (the Lyapunov time) into the future. Another example is Earth's axial tilt which, thanks to friction raised within Earth's mantle by tidal interactions with the Moon (see below) will be incomputable at some point between 1.5 and 4.5 billion years from now.

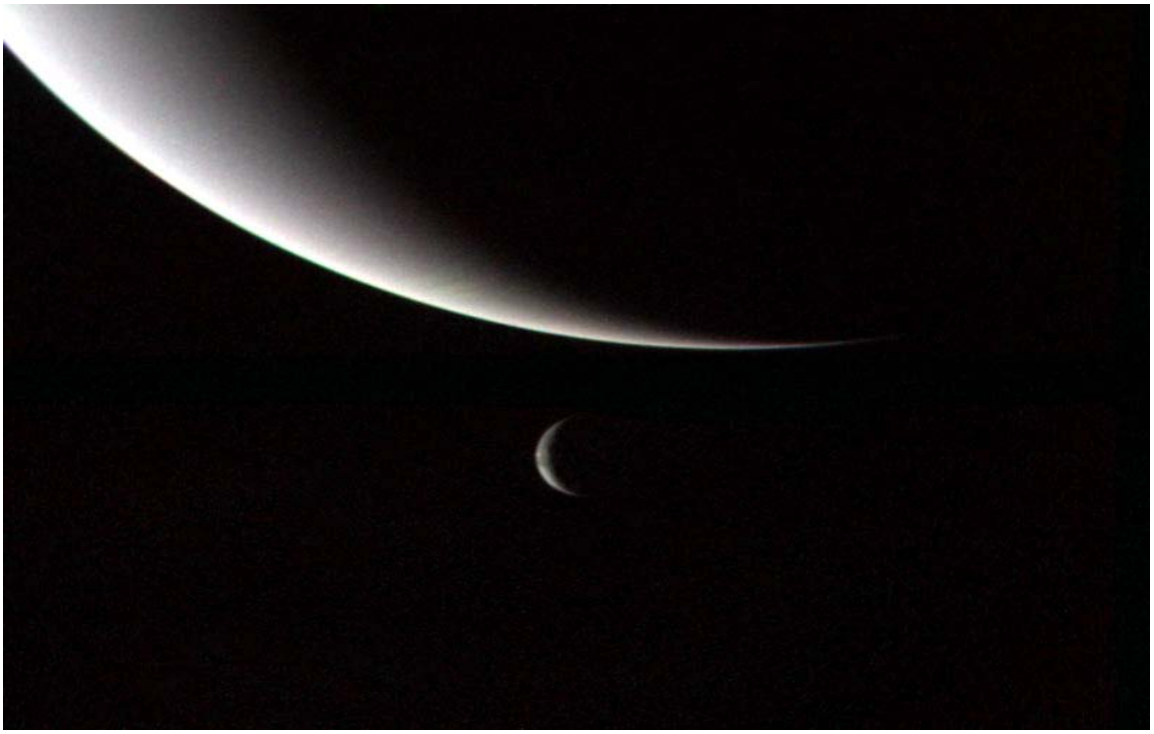
The outer planets' orbits are chaotic over longer timescales, such that they possess a Lyapunov time in the range of 2–230 million years. In all cases this means that the position of a planet along its orbit ultimately becomes impossible to predict with any certainty (so, for example, the timing of winter and summer become uncertain), but in some cases the orbits themselves may change dramatically. Such chaos manifests most strongly as changes in eccentricity, with some planets' orbits becoming significantly more—or less—elliptical.

Ultimately, the Solar System is stable in that none of the planets will collide with each other or be ejected from the system in the next few billion years. Beyond this, within five billion years or so Mars's eccentricity may grow to around 0.2, such that it lies on an Earth-crossing orbit, leading to a potential collision. In the same timescale, Mercury's eccentricity may grow even further, and a close encounter with Venus could theoretically eject it from the Solar System altogether or send it on a collision course with Venus or Earth.

Moon-ring systems

The evolution of moon systems is driven by tidal forces. A moon will raise a tidal bulge in the object it orbits (the primary) due to the differential gravitational force across diameter of the primary. If a moon is revolving in the same direction as the planet's rotation and the planet is rotating faster than the orbital period of the moon, the bulge will constantly be pulled ahead of the moon. In this situation, angular momentum is transferred from the rotation of the primary to the revolution of the satellite. The moon gains energy and gradually spirals outward, while the primary rotates more slowly over time.

The Earth and its Moon are one example of this configuration. Today, the Moon is tidally locked to the Earth; one of its revolutions around the Earth (currently about 29 days) is equal to one of its rotations about its axis, so it always shows one face to the Earth. The Moon will continue to recede from Earth, and Earth's spin will continue to slow gradually. In about 50 billion years, if they survive the Sun's expansion, the Earth and Moon will become tidally locked to each other; each will be caught up in what is called a "spin–orbit resonance" in which the Moon will circle the Earth in about 47 days and both Moon and Earth will rotate around their axes in the same time, each only visible from one hemisphere of the other. Other examples are the Galilean moons of Jupiter (as well as many of Jupiter's smaller moons) and most of the larger moons of Saturn.



Neptune and its moon Triton, taken by *Voyager 2*. Triton's orbit will eventually take it within Neptune's Roche limit, tearing it apart and possibly forming a new ring system.

A different scenario occurs when the moon is either revolving around the primary faster than the primary rotates, or is revolving in the direction opposite the planet's rotation. In these cases, the tidal bulge lags behind the moon in its orbit. In the former case, the direction of angular momentum transfer is reversed, so the rotation of the primary speeds up while the satellite's orbit shrinks. In the latter case, the angular momentum of the rotation and revolution have opposite signs, so transfer leads to decreases in the magnitude of each (that cancel each other out). In both cases, tidal deceleration causes the moon to spiral in towards the primary until it either is torn apart by tidal stresses, potentially creating a planetary ring system, or crashes into the planet's surface or atmosphere. Such a fate awaits the moons Phobos of Mars (within 30 to 50 million years), Triton of Neptune (in 3.6 billion years), Metis and Adrastea of Jupiter, and at least 16 small satellites of Uranus and Neptune. Uranus' Desdemona may even collide with one of its neighboring moons.

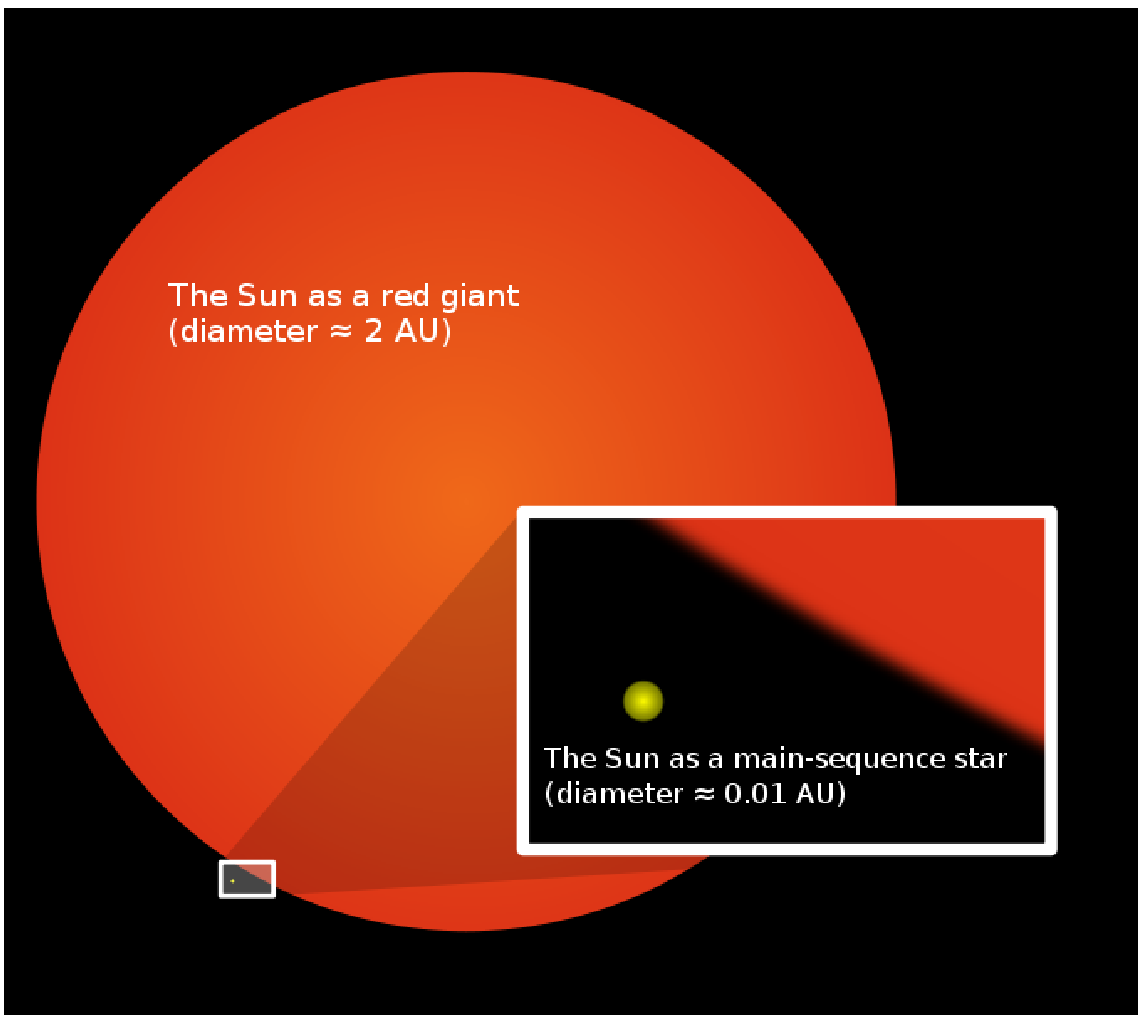
A third possibility is where the primary and moon are tidally locked to each other. In that case, the tidal bulge stays directly under the moon, there is no transfer of angular momentum, and the orbital period will not change. Pluto and Charon are an example of this type of configuration.

Prior to the 2004 arrival of the *Cassini–Huygens* spacecraft, the rings of Saturn were widely thought to be much younger than the Solar System and were not expected to survive beyond another 300 million years. Gravitational interactions with Saturn's moons were expected to gradually sweep the rings' outer edge toward the planet, with abrasion by meteorites and Saturn's gravity eventually taking the rest, leaving Saturn unadorned. However, data from the *Cassini* mission led scientists to revise that early view. Observations revealed 10 km-wide icy clumps of material that repeatedly break apart and reform, keeping the rings fresh. Saturn's rings are far more massive than the rings of the other gas giants. This large mass is believed to have preserved

Saturn's rings since the planet first formed 4.5 billion years ago, and is likely to preserve them for billions of years to come.

The Sun and planetary environments

In the long term, the greatest changes in the Solar System will come from changes in the Sun itself as it ages. As the Sun burns through its supply of hydrogen fuel, it gets hotter and burns the remaining fuel even faster. As a result, the Sun is growing brighter at a rate of ten percent every 1.1 billion years. In one billion years' time, as the Sun's radiation output increases, its circumstellar habitable zone will move outwards, making the Earth's surface hot enough that liquid water can no longer exist there naturally. At this point, all life on land will become extinct. Evaporation of water, a potent greenhouse gas, from the oceans' surface could accelerate temperature increase, potentially ending all life on Earth even sooner. During this time, it is possible that as Mars's surface temperature gradually rises, carbon dioxide and water currently frozen under the surface soil will release into the atmosphere, creating a greenhouse effect that will heat the planet until it achieves conditions parallel to Earth today, providing a potential future abode for life. By 3.5 billion years from now, Earth's surface conditions will be similar to those of Venus today.

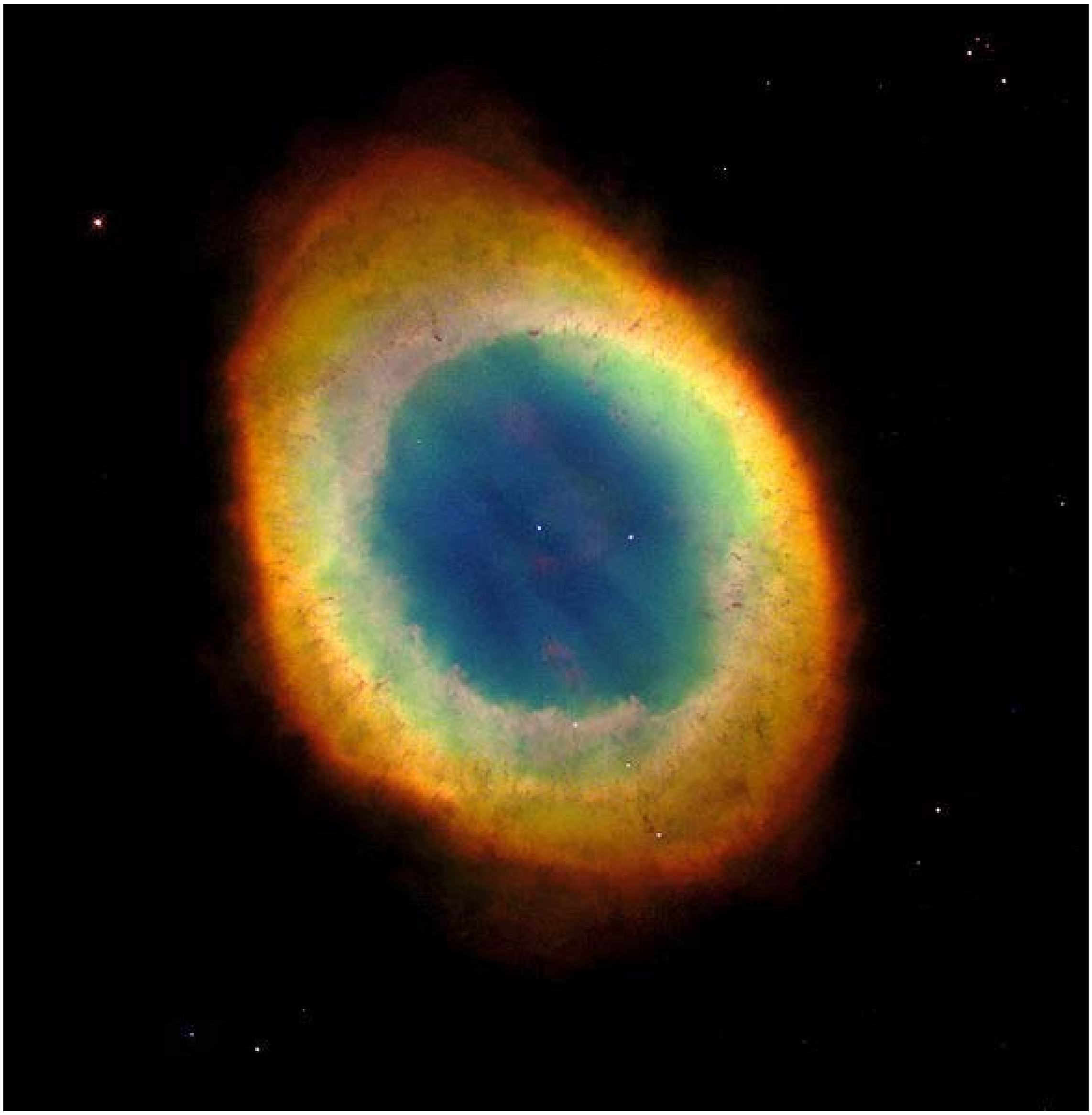


Relative size of our Sun as it is now (inset) compared to its estimated future size as a red giant

Around 5.4 billion years from now, the core of the Sun will become hot enough to trigger hydrogen fusion in its surrounding shell. This will cause the outer layers of the star to expand greatly, and the star will enter a phase of its life in which it is called a red giant. Within 7.5 billion years, the Sun will have expanded to a radius of 1.2 AU—256 times its current size. At the tip of the red giant branch, as a result of the vastly increased surface area, the Sun's surface will be much cooler (about 2600 K) than now and its luminosity much higher—up to 2,700 current solar luminosities. For part of its red giant life, the Sun will have a strong stellar wind that will carry away around 33% of its mass. During these times, it is possible that Saturn's moon Titan could achieve surface temperatures necessary to support life.

As the Sun expands, it will swallow the planets Mercury and, most likely, Venus. Earth's fate is less clear; although the Sun will envelop Earth's current orbit, the star's loss of mass (and thus weaker gravity) will cause the planets' orbits to move farther out. If it were only for this, Venus and Earth would probably escape incineration, but a 2008 study suggests that Earth will likely be swallowed up as a result of tidal interactions with the Sun's weakly bound outer envelope.

Gradually, the hydrogen burning in the shell around the solar core will increase the mass of the core until it reaches about 45% of the present solar mass. At this point the density and temperature will become so high that the fusion of helium into carbon will begin, leading to a helium flash; the Sun will shrink from around 250 to 11 times its present (main sequence) radius. Consequently, its luminosity will decrease from around 3,000 to 54 times its current level, and its surface temperature will increase to about 4770 K. The Sun will become a horizontal branch star, burning helium in its core in a stable fashion much like it burns hydrogen today. The helium-fusing stage will last only 100 million years. Eventually, it will have to again resort to the reserves of hydrogen and helium in its outer layers and will expand a second time, turning into what is known as an asymptotic giant branch star. Here the luminosity of the Sun will increase again, reaching about 2,090 present luminosities, and it will cool to about 3500 K. This phase lasts about 30 million years, after which, over the course of a further 100,000 years, the Sun's remaining outer layers will fall away, ejecting a vast stream of matter into space and forming a halo known (misleadingly) as a planetary nebula. The ejected material will contain the helium and carbon produced by the Sun's nuclear reactions, continuing the enrichment of the interstellar medium with heavy elements for future generations of stars.



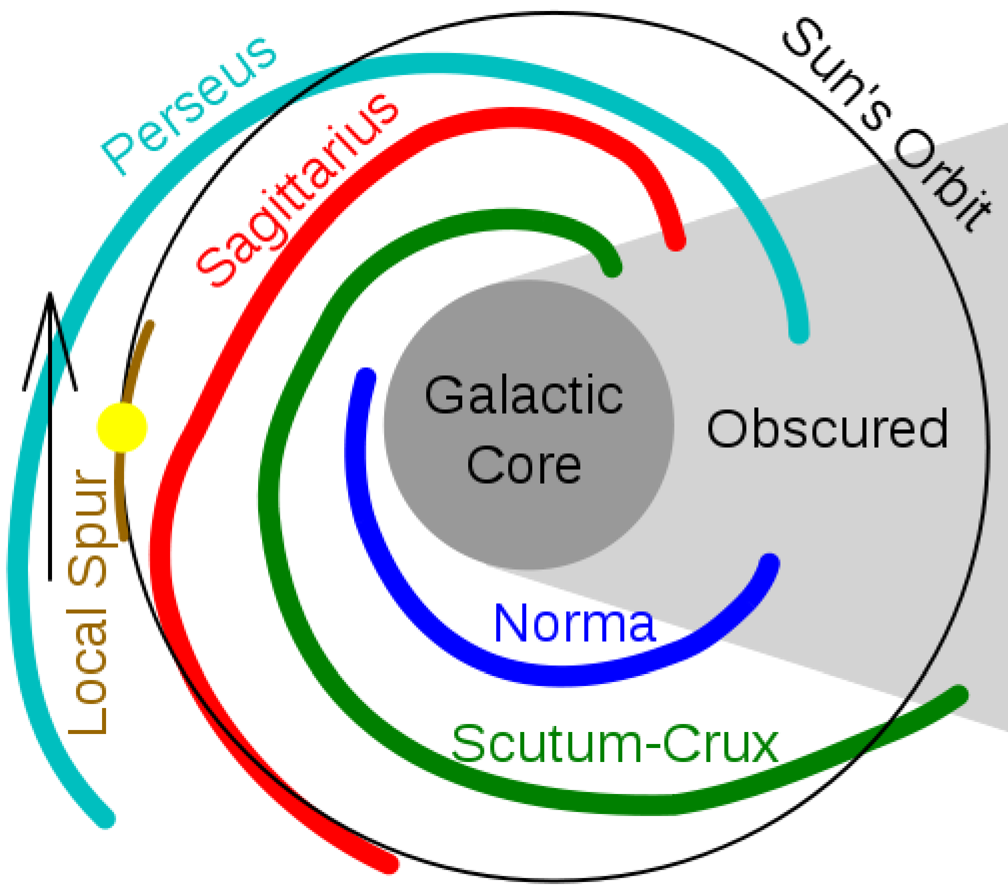
The Ring nebula, a planetary nebula similar to what the Sun will become

This is a relatively peaceful event, nothing akin to a supernova, which our Sun is too small to undergo as part of its evolution. Any observer present to witness this occurrence would see a massive increase in the speed of the solar wind, but not enough to destroy a planet completely. However, the star's loss of mass could send the orbits of the surviving planets into chaos, causing some to collide, others to be ejected from the Solar System, and still others to be torn apart by tidal interactions. Afterwards, all that will remain of the Sun is a white dwarf, an extraordinarily dense object, 54% its original mass but only the size of the Earth. Initially, this white dwarf may be 100 times as luminous as the Sun is now. It will consist entirely of degenerate carbon and oxygen, but will never reach temperatures hot enough to fuse these elements. Thus the white dwarf Sun will gradually cool, growing dimmer and dimmer.

As the Sun dies, its gravitational pull on the orbiting bodies such as planets, comets and asteroids will weaken due to its mass loss. All remaining planets' orbits will expand; if Venus, Earth, and

Mars still exist, their orbits will lie roughly at 1.4 AU (210,000,000 km), 1.9 AU (280,000,000 km), and 2.8 AU (420,000,000 km). They and the other remaining planets will become dark, frigid hulks, completely devoid of any form of life. They will continue to orbit their star, their speed slowed due to their increased distance from the Sun and the Sun's reduced gravity. Two billion years later, when the Sun has cooled to the 6000–8000K range, the carbon and oxygen in the Sun's core will freeze, with over 90% of its remaining mass assuming a crystalline structure. Eventually, after trillions more years, the Sun will finally cease to shine altogether, becoming a black dwarf.

Galactic interaction



Location of the Solar System within our galaxy

The Solar System travels alone through the Milky Way galaxy in a circular orbit approximately 30,000 light years from the galactic centre. Its speed is about 220 km/s. The period required for the Solar System to complete one revolution around the galactic centre, the galactic year, is in the range of 220–250 million years. Since its formation, the Solar System has completed at least 20 such revolutions.

A number of scientists have speculated that the Solar System's path through the galaxy is a factor in the periodicity of mass extinctions observed in the Earth's fossil record. One hypothesis supposes that vertical oscillations made by the Sun as it orbits the galactic centre cause it to regularly pass through the galactic plane. When the Sun's orbit takes it outside the galactic disc, the influence of the galactic tide is weaker; as it re-enters the galactic disc, as it does every 20–25 million years, it comes under the influence of the far stronger "disc tides", which, according to mathematical models, increase the flux of Oort cloud comets into the Solar System by a factor of 4, leading to a massive increase in the likelihood of a devastating impact.

However, others argue that the Sun is currently close to the galactic plane, and yet the last great extinction event was 15 million years ago. Therefore the Sun's vertical position cannot alone explain such periodic extinctions, and that extinctions instead occur when the Sun passes through the galaxy's spiral arms. Spiral arms are home not only to larger numbers of molecular clouds, whose gravity may distort the Oort cloud, but also to higher concentrations of bright blue giant stars, which live for relatively short periods and then explode violently as supernovae.

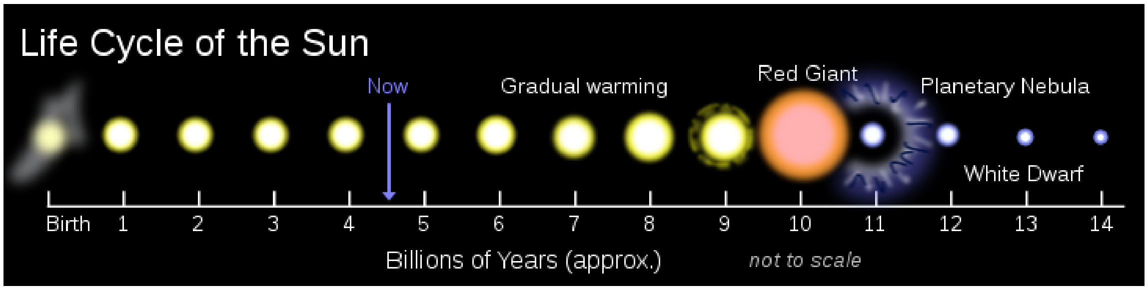
Galactic collision and planetary disruption

Although the vast majority of galaxies in the Universe are moving away from the Milky Way, the Andromeda Galaxy, the largest member of our Local Group of galaxies, is heading towards it at about 120 km/s. In 2 billion years, Andromeda and the Milky Way will collide, causing both to deform as tidal forces distort their outer arms into vast tidal tails. If this initial disruption occurs, astronomers calculate a 12% chance that the Solar System will be pulled outward into the Milky Way's tidal tail and a 3% chance that it will become gravitationally bound to Andromeda and thus a part of that galaxy. After a further series of glancing blows, during which the likelihood of the Solar System's ejection rises to 30%, the galaxies' supermassive black holes will merge. Eventually, in roughly 7 billion years, the Milky Way and Andromeda will complete their merger into a giant elliptical galaxy. During the merger, if there is enough gas, the increased gravity will force the gas to the centre of the forming elliptical galaxy. This may lead to a short period of intensive star formation called a starburst. In addition the infalling gas will feed the newly formed black hole transforming it into an active galactic nucleus. The force of these interactions will likely push the Solar System into the new galaxy's outer halo, leaving it relatively unscathed by the radiation from these collisions.

It is a common misconception that this collision will disrupt the orbits of the planets in the Solar System. While it is true that the gravity of passing stars can detach planets into interstellar space, distances between stars are so great that the likelihood of the Milky Way-Andromeda collision causing such disruption to any individual star system is negligible. While the Solar System as a whole could be affected by these events, the Sun and planets are not expected to be disturbed.

However, over time, the cumulative probability of a chance encounter with a star increases, and disruption of the planets becomes all but inevitable. Assuming that the Big Crunch or Big Rip scenarios for the end of the universe do not occur, calculations suggest that the gravity of passing stars will have completely stripped the dead Sun of its remaining planets within 1 quadrillion (10^{15}) years. This point marks the end of the Solar System. While the Sun and planets may survive, the Solar System, in any meaningful sense, will cease to exist.

Chronology



The time frame of the Solar System's formation has been determined using radiometric dating. Scientists estimate that the Solar System is 4.6 billion years old. The oldest known mineral grains on Earth are approximately 4.4 billion years old. Rocks this old are rare, as Earth's surface is constantly being reshaped by erosion, volcanism, and plate tectonics. To estimate the age of the Solar System, scientists use meteorites, which were formed during the early condensation of the solar nebula. Almost all meteorites are found to have an age of 4.6 billion years, suggesting that the Solar System must be at least this old.

Studies of discs around other stars have also done much to establish a time frame for Solar System formation. Stars between one and three million years old possess discs rich in gas, whereas discs around stars more than 10 million years old have little to no gas, suggesting that gas giant planets within them have ceased forming.

Timeline of Solar System evolution

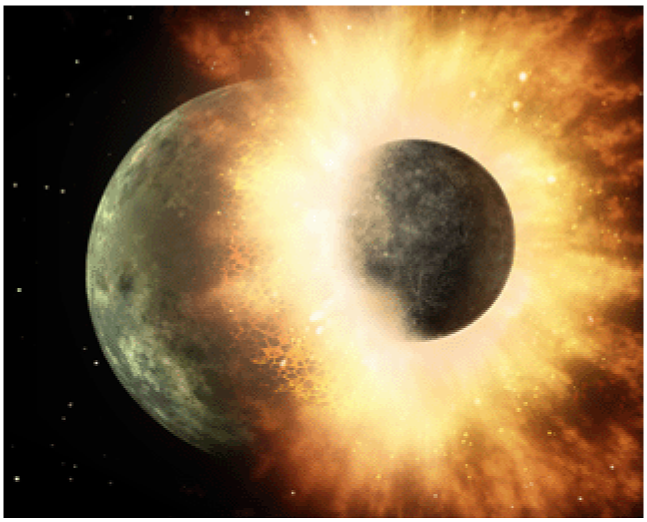
Note: All dates and times in this chronology are approximate and should be taken as an order of magnitude indicator only.

Phase	Time since formation of the Sun	Event
Pre-Solar System	Billions of years before the formation of the Solar System	Previous generations of stars live and die, injecting heavy elements into the interstellar medium out of which the Solar System formed.
	~ 50 million years before formation of the Solar System	If the Solar System formed in an Orion nebula-like star-forming region, the most massive stars are formed, live their lives, die, and explode in supernovae. One supernova possibly triggers the formation of the Solar System.
Formation of Sun	**0**–100,000 years	Pre-solar nebula forms and begins to collapse. Sun begins to form.
	100,000 – 50 million years	Sun is a T Tauri protostar.
	100,000 - 10 million years	Outer planets form. By 10 million years, gas in the protoplanetary disc has been blown away, and outer planet formation is likely complete.

	10 million - 100 million years	Terrestrial planets and the Moon form. Giant impacts occur. Water delivered to Earth.
Main sequence	50 million years	Sun becomes a main sequence star.
	200 million years	Oldest known rocks on the Earth formed.
	500 million – 600 million years	Resonance in Jupiter and Saturn's orbits moves Neptune out into the Kuiper belt. Late Heavy Bombardment occurs in the inner Solar System.
	800 million years	Oldest known life on Earth. Oort cloud reaches maximum mass.
	4.6 billion years	**Today**. Sun remains a main sequence star, continually growing warmer and brighter by ~10% every 1 billion years.
	6 billion years	Sun's habitable zone moves outside of the Earth's orbit, possibly shifting onto Mars' orbit.
	7 billion years	The Milky Way and Andromeda Galaxy begin to collide. Slight chance the Solar System could be captured by Andromeda before the two galaxies fuse completely.
Post-main sequence	10 billion – 12 billion years	Sun starts burning hydrogen in a shell surrounding its core, ending its main sequence life. Sun begins to ascend the red giant branch of the Hertzsprung-Russell diagram, growing dramatically more luminous (by a factor of up to 2,700), larger (by a factor of up to 250 in radius), and cooler (down to 2600 K): Sun is now a red giant. Mercury and possibly Venus and Earth are swallowed. Saturn's moon Titan may become habitable.
	~ 12 billion years	Sun passes through helium-burning horizontal branch and asymptotic giant branch phases, losing a total of ~30% of its mass in all post-main sequence phases. Asymptotic giant branch phase ends with the ejection of a planetary nebula, leaving the core of the Sun behind as a white dwarf.
Remnant Sun	> 12 billion years	The white dwarf Sun, no longer producing energy, begins to cool and dim continuously, eventually reaching a black dwarf state.
	~ 1 quadrillion years (10^{15} years)	Sun cools to 5 K. Gravity of passing stars detaches planets from orbits. Solar System ceases to exist.

Chapter- 3

Giant Impact Hypothesis



Artist's depiction of the giant impact that is hypothesized to have formed the Moon

The **giant impact hypothesis** proposes that the Moon was created out of the debris left over from a collision between the young Earth and a Mars-sized body. This is the favored scientific hypothesis for the formation of the Moon. Evidence for this hypothesis includes Moon samples which indicate the surface of the Moon was once molten, the Moon's apparently relatively small iron core and a lower density than the Earth, and evidence of similar collisions in other star systems (which result in debris disks). The colliding body is sometimes called Theia (or Orpheus) for the mythical Greek Titan who was the mother of Selene, the goddess of the moon.

There remain several unanswered issues surrounding this hypothesis. Lunar oxygen isotopic ratios are essentially identical to Earth's, with no evidence of a contribution from another solar body. Also, lunar samples do not have expected ratios of volatile elements, iron oxide, or siderophilic elements, and there is no evidence to suggest that the Earth ever had the magma ocean implied by this hypothesis.

Origins

In 1898, George Howard Darwin made an early suggestion that the Earth and Moon had once been one body. Darwin's hypothesis was that a molten Moon had been spun from the Earth because of centrifugal forces, and this became the dominant academic explanation. Using

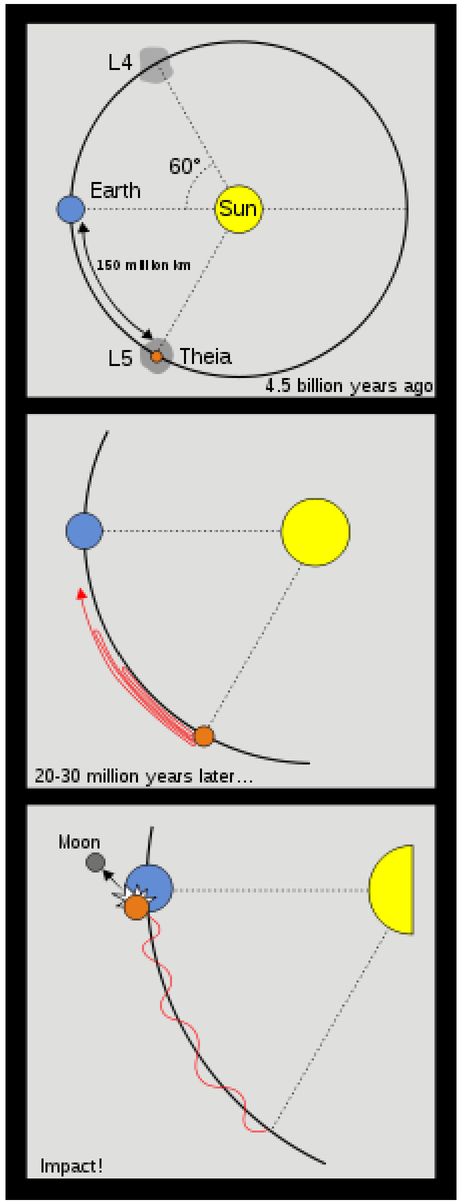
Newtonian mechanics, he calculated that the Moon had actually orbited much closer in the past and was drifting away from the Earth. This drifting was later confirmed by American and Soviet experiments using laser ranging targets placed on the Moon.

However, Darwin's calculations could not resolve the mechanics required to trace the Moon backwards to the surface of the Earth. In 1946, Reginald Aldworth Daly of Harvard University challenged Darwin's explanation, adjusting it to postulate that the creation of the Moon was caused by an impact rather than centrifugal forces. Little attention was paid to Professor Daly's challenge until a conference on satellites in 1974 where it was reintroduced. It was then republished in *Icarus* in 1975 by Drs. William K. Hartmann and Donald R. Davis. Their models suggested that, at the end of the planet formation period, several satellite-sized bodies had formed that could collide with the planets or be captured. They proposed that one of these objects may have collided with the Earth, ejecting refractory, volatile-poor dust that could coalesce to form the Moon. This collision could help explain the unique geological properties of the Moon.

A similar approach was taken by Alfred G. W. Cameron and William Ward, who suggested that the Moon was formed by the tangential impact of a body the size of Mars. The outer silicates of the colliding body would mostly be vaporized, whereas a metallic core would not. Hence, most of the collisional material sent into orbit would consist of silicates, leaving the coalescing Moon deficient in iron. The more volatile materials that were emitted during the collision would likely escape the Solar System, whereas silicates would tend to coalesce.

Theia

The name of the hypothesized protoplanet is derived from the mythical Greek goddess Theia, a Titan who gave birth to the Moon goddess Selene. According to the giant impact hypothesis, Theia formed alongside the other planet size bodies in the Solar System about 4.6 Ga (4.6 billion years ago), and was approximately the size of Mars.



One suggested pathway for the **Big Splash** as viewed from the direction of the south pole

One formation theory is that Theia materialized at the L_4 or L_5 Lagrangian points relative to Earth (in about the same orbit and about 60° ahead or behind), similar to a trojan asteroid. The stability of Theia's orbit was affected when its growing mass exceeded a threshold of about 10% of the Earth's mass. Gravitational perturbations by planetesimals caused Theia to depart from its stable Lagrangian location, and subsequent interactions with proto-Earth caused the two bodies to collide.

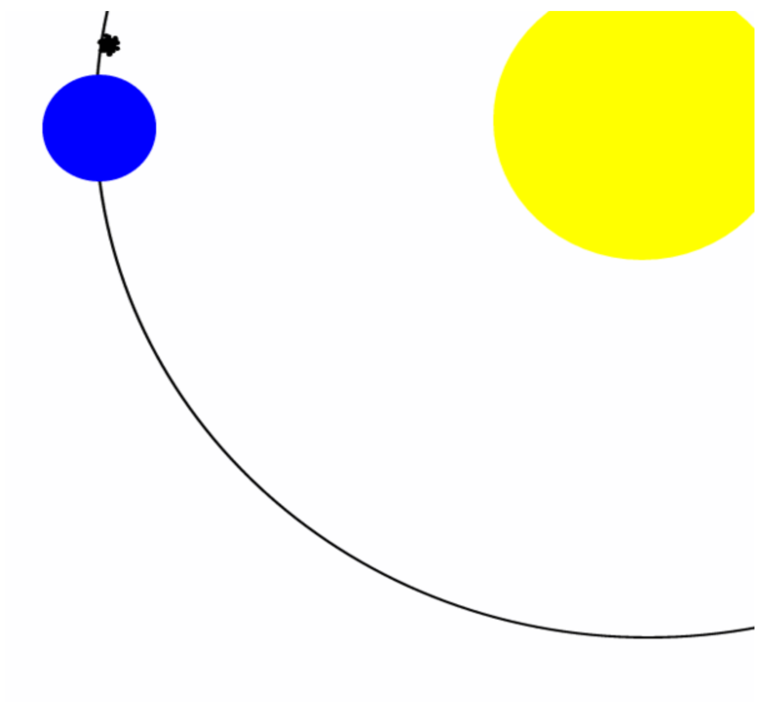
Astronomers think the collision between Earth and Theia happened about 4.53 Ga; about 30-50 million years after the rest of the Solar System formed. However, evidence presented in 2008 suggests that the collision may have occurred later, at about 4.48 Ga.

Impact

In astronomical terms, the impact would have been of moderate velocity. Theia is thought to have struck the Earth at an oblique angle when the planet was nearly fully formed. Computer simulations of this "late-impact" scenario suggest an impact angle of about 45° and an initial impactor velocity below 4 km/s. Theia's iron core sank into the young Earth's core, as most of Theia's mantle and a significant portion of the Earth's mantle and crust were ejected into orbit around the Earth. This material quickly coalesced into the Moon (possibly within less than a month, but in no more than a century). Estimates based on computer simulations of such an event suggest that some two percent of the original mass of Theia ended up as an orbiting ring of debris, and about half of this matter coalesced into the Moon. The Earth would have gained significant amounts of angular momentum and mass from such a collision. Regardless of the rotation and inclination the Earth had before the impact, it would have had a day some five hours long after the impact, and the Earth's equator would have shifted closer to the plane of the Moon's orbit.

It has been suggested that other significant objects may have been created by the impact, which could have remained in orbit between the Earth and Moon, stuck in Lagrangian points. Such objects may have stayed within the Earth-Moon system for up to 100 million years, until the gravitational tugs of other planets destabilized the system enough to free the objects.

Evidence



Theia possibly forming in Earth's L_5 point and then drifting into impact. The progresses in one-year steps (before impact) making Earth appear not to move. The view is of the south pole.

Indirect evidence for this impact scenario comes from rocks collected during the Apollo Moon landings, which show oxygen isotope ratios identical to those of Earth. The highly anorthositic composition of the lunar crust, as well as the existence of KREEP-rich samples, gave rise to the idea that a large portion of the Moon was once molten, and a giant impact scenario could easily have supplied the energy needed to form such a magma ocean. Several lines of evidence show that if the Moon has an iron-rich core, it must be small. In particular, the mean density, moment of inertia, rotational signature, and magnetic induction response all suggest that the radius of the core is less than about 25% the radius of the Moon, in contrast to about 50% for most of the other terrestrial bodies. Impact conditions can be found that give rise to a Moon that formed mostly from the mantles of the Earth and impactor, with the core of the impactor accreting to the Earth, and which satisfy the angular momentum constraints of the Earth-Moon system.

Warm silica-rich dust and abundant SiO gas, products of high velocity (> 10 km/sec) impacts between rocky bodies has been detected around the nearby (29 pc distant) young (~12 My old) Beta Pic Moving Group star HD172555 by the Spitzer Space Telescope. A belt of warm dust in a zone between 0.25AU and 2AU from the young star HD 23514 in the Pleiades cluster appears similar to the predicted results of Theia's collision with the embryonic Earth, and has been interpreted as the result of planet-sized objects colliding with each other. This is similar to another belt of warm dust detected around the star BD +20°307 (HIP 8920, SAO 75016).

Difficulties

This lunar origin hypothesis has some difficulties which have yet to be resolved. These difficulties include:

- The ratios of the Moon's volatile elements are not explained by the giant impact hypothesis. If the giant impact hypothesis is correct, they must be due to some other cause.
- There is no evidence that the Earth ever had a magma ocean (an implied result of the giant impact hypothesis), and it is likely there exists material which has never been processed by a magma ocean.
- Iron oxide (FeO) content of 13% of the bulk Moon properties rule out the derivation of the proto-lunar material from any but a small fraction of Earth's mantle
- If the bulk of the proto-lunar material had come from the impactor, the Moon should be enriched in siderophilic elements, when it is actually deficient in those.
- The presence of volatiles such as water trapped in lunar basalts is more difficult to explain if the impact caused a catastrophic heating event.
- The Moon's oxygen isotopic ratios are essentially identical to those of Earth. Oxygen isotopic ratios, which can be measured very precisely, yield a unique and distinct signature for each solar system body. If Theia was a separate proto-planet, it would have likely had a different oxygen isotopic signature than Earth, as would the ejected mixed material.

Alternate hypotheses

Other mechanisms which have been suggested at various times for the Moon's origin are that the Moon was spun off of the Earth's molten, blobular surface by centrifugal force, that it was formed elsewhere and later captured by the Earth's gravitational field, and that the Moon formed at the same time and place as the Earth from the same accretion disk. Each of these hypotheses is claimed to lack a mechanism to account for the high angular momentum of the Earth–Moon system.

Chapter- 4

Origin and Evolution of the Moon

Moon ☾

A moon just past full as seen from Earth's northern hemisphere

Designations	
Adjective	lunar
Orbital characteristics	

Perigee	363,104 km (0.0024 AU)
Apogee	405,696 km (0.0027 AU)
Semi-major axis	384,399 km (0.00257 AU)
Eccentricity	0.0549
Orbital period	27.321582 d (27 d 7 h 43.1 min)
Synodic period	29.530589 d (29 d 12 h 44 min 2.9 s)
Average orbital speed	1.022 km/s
Inclination	5.145° to the ecliptic (between 18.29° and 28.58° to Earth's equator)
Longitude of ascending node	regressing by one revolution in 18.6 years
Argument of perigee	progressing by one revolution in 8.85 years
Satellite of	Earth
Physical characteristics	
Mean radius	1,737.10 km (0.273 Earths)
Equatorial radius	1,738.14 km (0.273 Earths)
Polar radius	1,735.97 km (0.273 Earths)
Flattening	0.00125
Circumference	10,921 km (equatorial)
Surface area	3.793×10^{7} km^2 (0.074 Earths)
Volume	2.1958×10^{10} km^3 (0.020 Earths)
Mass	7.3477×10^{22} kg (0.0123 Earths)
Mean density	3.3464 g/cm^3

Equatorial surface gravity	1.622 m/s^2 (0.165 4 g)
Escape velocity	2.38 km/s
Sidereal rotation period	27.321582 d (synchronous)
Equatorial rotation velocity	4.627 m/s
Axial tilt	1.5424° (to ecliptic) 6.687° (to orbit plane)
Albedo	0.136

Surface temp.	**min**	**mean**	**max**
equator	100 K	220 K	390 K
85°N	70 K	130 K	230 K

Apparent magnitude	−2.5 to −12.9 −12.74 (mean full Moon)
Angular diameter	29.3 to 34.1 arcminutes
Atmosphere	
Surface pressure	10^{-7} Pa (day) 10^{-10} Pa (night)
Composition	Ar, He, Na, K, H, Rn

The **Moon** is Earth's only natural satellite and is the fifth largest satellite in the Solar System. It is the largest natural satellite in the Solar System relative to the size of its planet, a quarter the diameter of Earth and 1/81 its mass, and is the second densest satellite after Io. It is in synchronous rotation with Earth, always showing the same face; the near side is marked with dark volcanic maria among the bright ancient crustal highlands and prominent impact craters. It is the brightest object in the sky after the Sun, although its surface is actually very dark, with a similar reflectance to coal. Its prominence in the sky and its regular cycle of phases have since ancient times made the Moon an important cultural influence on language, the calendar, art and mythology. The Moon's gravitational influence produces the ocean tides and the minute lengthening of the day. The Moon's current orbital distance, about thirty times the diameter of the Earth, causes it to be the same size in the sky as the Sun—allowing the Moon to cover the Sun precisely in total solar eclipses.

The Moon is the only celestial body on which humans have made a manned landing. While the Soviet Union's Luna programme was the first to reach the Moon with unmanned spacecraft, the United States' NASA Apollo program achieved the only manned missions to date, beginning with the first manned lunar orbiting mission by Apollo 8 in 1968, and six manned lunar landings between 1969 and 1972—the first being Apollo 11 in 1969. These missions returned over 380 kg of lunar rocks, which have been used to develop a detailed geological understanding of the Moon's origins (it is thought to have formed some 4.5 billion years ago in a giant impact), the formation of its internal structure, and its subsequent history.

Since the Apollo 17 mission in 1972, the Moon has been visited only by unmanned spacecraft, notably by Soviet Lunokhod rovers. Since 2004, Japan, China, India, the United States, and the European Space Agency have each sent lunar orbiters. These spacecraft have contributed to confirming the discovery of lunar water ice in permanently shadowed craters at the poles and bound into the lunar regolith. Future manned missions to the Moon are planned but not yet underway; the Moon remains, under the Outer Space Treaty, free to all nations to explore for peaceful purposes.

Name and etymology

The English proper name for Earth's natural satellite is "the Moon". The noun *moon* derives from *moone* (around 1380), which developed from *mone* (1135), which derives from Old English *mōna* (dating from before 725), which, like all Germanic language cognates, ultimately stems from Proto-Germanic **mǣnōn*.

The principal modern English adjective pertaining to the Moon is *lunar*, derived from the Latin *Luna*. Another less common adjective is *selenic*, derived from the Ancient Greek *Selene* (*Σελήνη*), from which the prefix "seleno-" (as in *selenography*) is derived.

Formation

Several mechanisms have been proposed for the Moon's formation 4.527 ± 0.010 billion years ago, some 30–50 million years after the origin of the Solar System. These include the fission of the Moon from the Earth's crust through centrifugal forces, which would require too great an initial spin of the Earth, the gravitational capture of a pre-formed Moon, which would require an unfeasibly extended atmosphere of the Earth to dissipate the energy of the passing Moon, and the co-formation of the Earth and the Moon together in the primordial accretion disk, which does not explain the depletion of metallic iron in the Moon. These hypotheses also cannot account for the high angular momentum of the Earth–Moon system.

The prevailing hypothesis today is that the Earth–Moon system formed as a result of a giant impact: a Mars-sized body hit the nearly formed proto-Earth, blasting material into orbit around the proto-Earth, which accreted to form the Moon. Giant impacts are thought to have been common in the early Solar System. Computer simulations modelling a giant impact are consistent with measurements of the angular momentum of the Earth–Moon system, and the small size of the lunar core; they also show that most of the Moon came from the impactor, not from the proto-Earth. However, meteorites show that other inner Solar System bodies such as Mars and Vesta have very different oxygen and tungsten isotopic compositions to the Earth, while the Earth and Moon have near-identical isotopic compositions. Post-impact mixing of the vaporized material between the forming Earth and Moon could have equalized their isotopic compositions, although this is debated.

The large amount of energy released in the giant impact event and the subsequent reaccretion of material in Earth orbit would have melted the outer shell of the Earth, forming a magma ocean. The newly formed Moon would also have had its own lunar magma ocean; estimates for its depth range from about 500 km to the entire radius of the Moon.

Physical characteristics

Internal structure

Chemical composition of the lunar surface regolith (derived from crustal rocks)

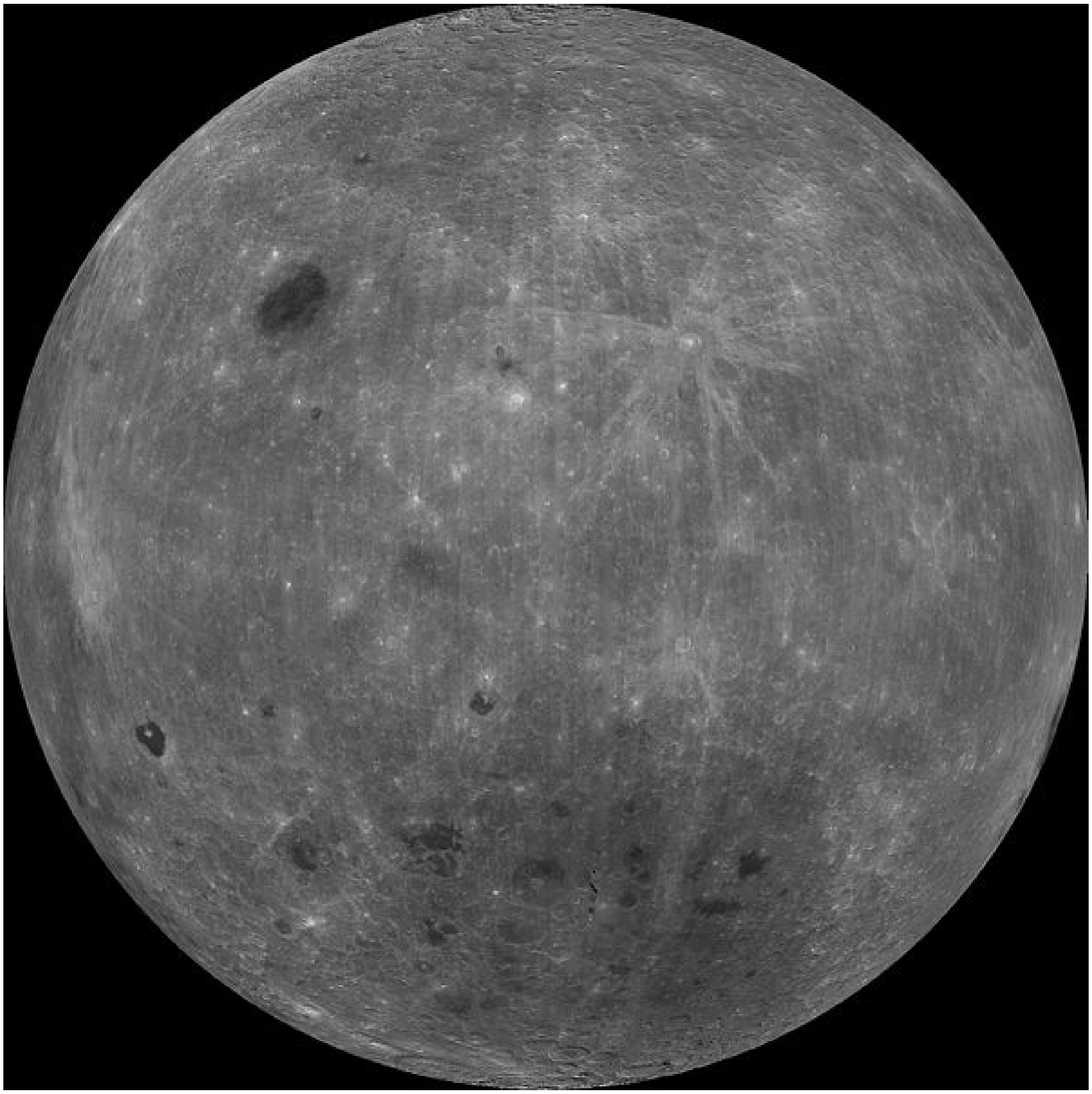
Compound	Formula	Composition (wt %)	
		Maria	Highlands
silica	SiO_2	45.4%	45.5%
alumina	Al_2O_3	14.9%	24.0%
lime	CaO	11.8%	15.9%
iron(II) oxide	FeO	14.1%	5.9%
magnesia	MgO	9.2%	7.5%
titanium dioxide	TiO_2	3.9%	0.6%
sodium oxide	Na_2O	0.6%	0.6%
Total		**99.9%**	**100.0%**

The Moon is a differentiated body: it has a geochemically distinct crust, mantle, and core. This structure is thought to have developed through the fractional crystallization of a global magma ocean shortly after the Moon's formation 4.5 billion years ago. Crystallization of this magma ocean would have created a mafic mantle from the precipitation and sinking of the minerals olivine, clinopyroxene, and orthopyroxene; after about three-quarters of the magma ocean had crystallised, lower-density plagioclase minerals could form and float into a crust on top. The final liquids to crystallise would have been initially sandwiched between the crust and mantle, with a high abundance of incompatible and heat-producing elements. Consistent with this, geochemical mapping from orbit shows the crust is mostly anorthosite, and moon rock samples of the flood lavas erupted on the surface from partial melting in the mantle confirm the mafic mantle composition, which is more iron rich than that of Earth. Geophysical techniques suggest that the crust is on average ~50 km thick.

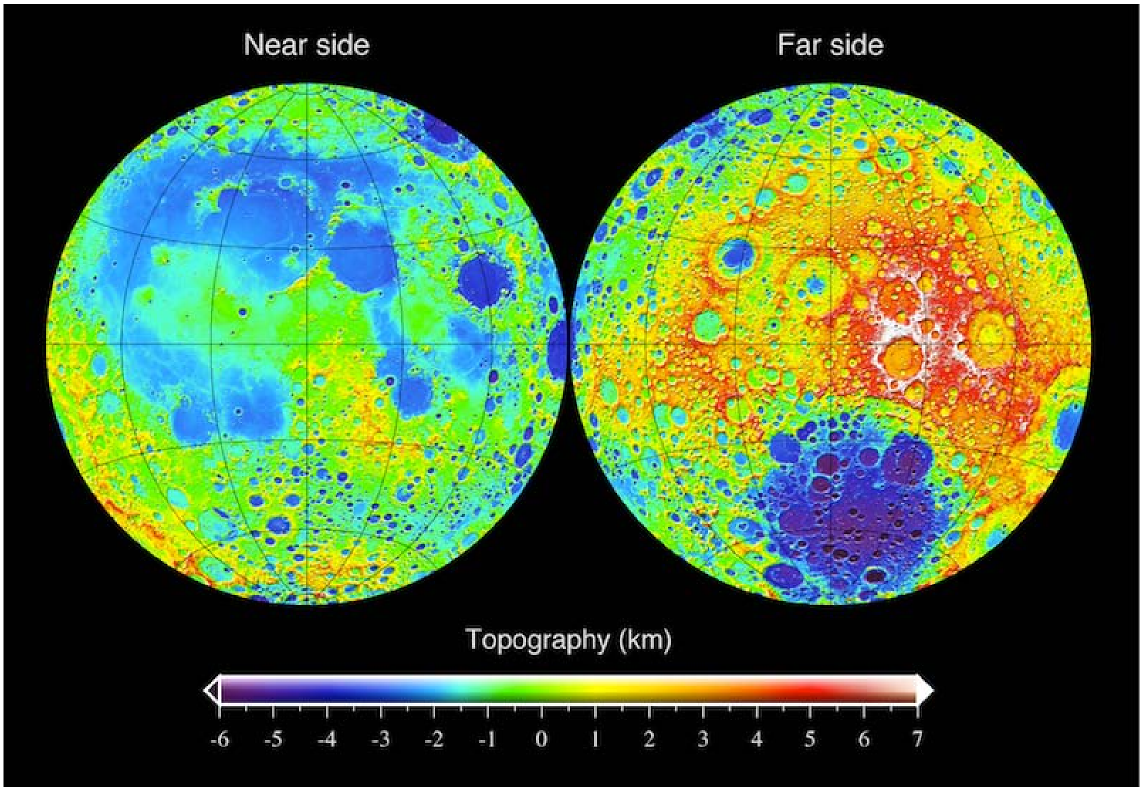
The Moon is the second densest satellite in the Solar System after Io. However, the core of the Moon is small, with a radius of about 350 km or less; this is only ~20% the size of the Moon, in contrast to the ~50% of most other terrestrial bodies. Its composition is not well constrained, but it is probably metallic iron alloyed with a small amount of sulphur and nickel; analyses of the Moon's time-variable rotation indicate that it is at least partly molten.

Surface geology

Near side of the Moon



Far side of the Moon. Note the lack of dark maria.



Topography of the Moon

The Moon is in synchronous rotation: it rotates about its axis in about the same time it takes to orbit the Earth. This results in it nearly always keeping the same face turned towards the Earth. The Moon used to rotate at a faster rate, but early in its history, its rotation slowed and became locked in this orientation as a result of frictional effects associated with tidal deformations caused by the Earth. The side of the Moon that faces Earth is called the near side, and the opposite side the far side. The far side is often called the "dark side," but in fact, it is illuminated as often as the near side: once per lunar day, during the new Moon phase we observe on Earth when the near side is dark.

The topography of the Moon has been measured with laser altimetry and stereo image analysis. The most visible topographic feature is the giant far side South Pole – Aitken basin, some 2,240 km in diameter, the largest crater on the Moon and the largest known crater in the Solar System. At 13 km deep, its floor is the lowest elevation on the Moon. The highest elevations are found just to its north-east, and it has been suggested that this area might have been thickened by the oblique formation impact of South Pole – Aitken. Other large impact basins, such as Imbrium, Serenitatis, Crisium, Smythii, and Orientale, also possess regionally low elevations and elevated rims. The lunar far side is on average about 1.9 km higher than the near side.

Volcanic features

The dark and relatively featureless lunar plains which can clearly be seen with the naked eye are called *maria* (Latin for "seas"; singular *mare*), since they were believed by ancient astronomers to be filled with water. They are now known to be vast solidified pools of ancient basaltic lava. While similar to terrestrial basalts, the mare basalts have much higher abundances of iron and are completely lacking in minerals altered by water. The majority of these lavas erupted or flowed into the depressions associated with impact basins. Several geologic provinces containing shield volcanoes and volcanic domes are found within the near side maria.

Maria are found almost exclusively on the near side of the Moon, covering 31% of the surface on the near side, compared with a few scattered patches on the far side covering only 2%. This is thought to be due to a concentration of heat-producing elements under the crust on the near side, seen on geochemical maps obtained by *Lunar Prospector'*s gamma-ray spectrometer, which would have caused the underlying mantle to heat up, partially melt, rise to the surface and erupt. Most of the Moon's mare basalts erupted during the Imbrian period, 3.0–3.5 billion years ago, although some radiometrically dated samples are as old as 4.2 billion years, and the youngest eruptions, dated by crater counting, appear to have been only 1.2 billion years ago.

The lighter-coloured regions of the Moon are called *terrae*, or more commonly *highlands*, since they are higher than most maria. They have been radiometrically dated as forming 4.4 billion years ago, and may represent plagioclase cumulates of the lunar magma ocean. In contrast to the Earth, no major lunar mountains are believed to have formed as a result of tectonic events.

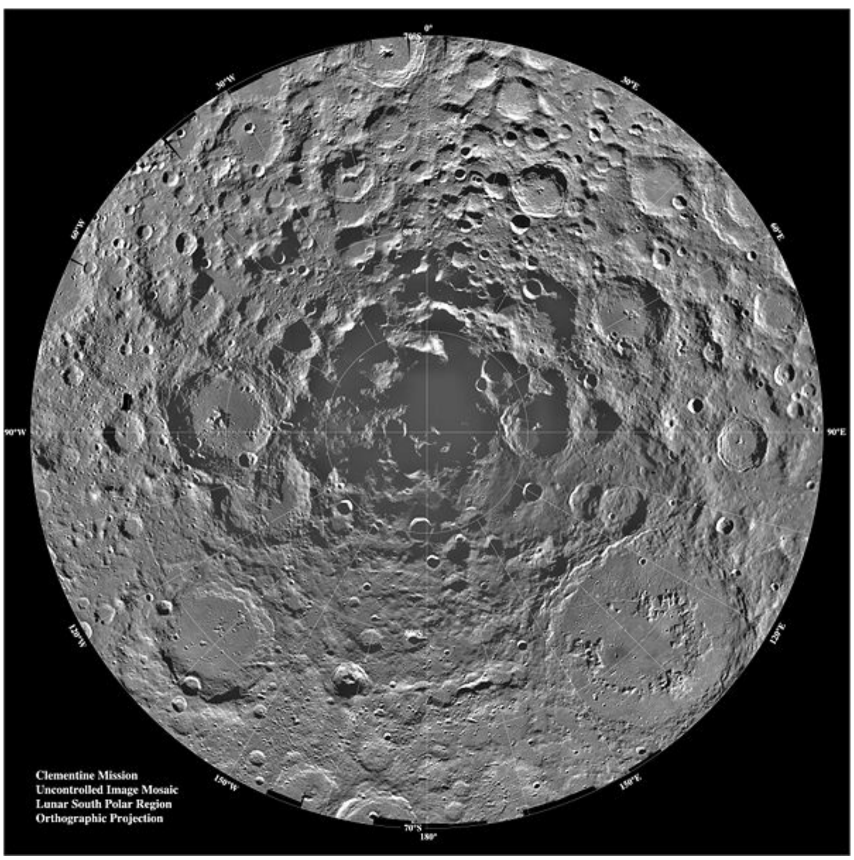
Impact craters

Lunar crater Daedalus on the Moon's far side

The other major geologic process that has affected the Moon's surface is impact cratering, with craters formed when asteroids and comets collide with the lunar surface. There are estimated to be roughly 300,000 craters wider than 1 km on the Moon's near side alone. These are named for scholars, scientists, artists and explorers. The lunar geologic timescale is based on the most prominent impact events, including Nectaris, Imbrium, and Orientale, structures characterized by multiple rings of uplifted material, typically hundreds to thousands of kilometres in diameter and associated with a broad apron of ejecta deposits that form a regional stratigraphic horizon. The lack of an atmosphere, weather and recent geological processes mean that many of these craters are well-preserved. While only a few multi-ring basins have been definitively dated, they are useful for assigning relative ages. Since impact craters accumulate at a nearly constant rate, counting the number of craters per unit area can be used to estimate the age of the surface. The radiometric ages of impact-melted rocks collected during the Apollo missions cluster between 3.8 and 4.1 billion years old: this has been used to propose a Late Heavy Bombardment of impacts.

Blanketed on top of the Moon's crust is a highly comminuted (broken into ever smaller particles) and impact gardened surface layer called regolith, formed by impact processes. The finer regolith, the lunar soil of silicon dioxide glass, has a texture like snow and smell like spent gunpowder. The regolith of older surfaces is generally thicker than for younger surfaces: it varies in thickness from 10–20 m in the highlands and 3–5 m in the maria. Beneath the finely comminuted regolith layer is the *megaregolith*, a layer of highly fractured bedrock many kilometres thick.

Presence of water



Mosaic image of the lunar south pole as taken by *Clementine*: note permanent polar shadow

Liquid water cannot persist at the Moon's surface, and water vapour quickly evaporates, breaks up through photodissociation due to sunlight, and is lost to space. However, scientists have thought since the 1960s that water ice, deposited by impacting comets or produced by the reaction of oxygen-rich lunar rocks and hydrogen in the solar wind, could survive in the cold, permanently shadowed craters at the Moon's poles. These craters have been in shadow for the past two billion years, and computer simulations suggest that up to 14,000 km^2 might be in permanent shadow. The presence of usable quantities of water on the Moon is an important factor in rendering lunar habitation cost-effective, since transporting it from Earth would be prohibitively expensive.

Many different signatures of lunar water have since been found. In 1994, *Clementine*'s bistatic radar experiment found indications of small, frozen pockets of water close to the surface (though later Arecibo radar observations suggested these might be rocks ejected from young impact craters); *Lunar Prospector*'s neutron spectrometer indicated in 1998 that high concentrations of hydrogen are present in the upper metre of the regolith near the polar regions; in 2008, new analysis found small amounts of water in the interior of volcanic lava beads brought to Earth by Apollo 15. In September 2009, *Chandrayaan-1*'s imaging spectrometer detected water and hydroxyl absorption lines in reflected sunlight, evidence of large quantities of water on the Moon's surface, possibly as high as 1,000 ppm. Weeks later, the *LCROSS* mission flew its 2300 kg impactor into a permanently shadowed polar crater, and detected at least 100 kg of water in the plume of ejected material.

Gravity and magnetic fields

The gravitational field of the Moon has been measured through tracking the Doppler shift of radio signals emitted by orbiting spacecraft. The main lunar gravity features are mascons, large positive gravitational anomalies associated with some of the giant impact basins, partly caused by the dense mare basaltic lava flows that fill these basins. These anomalies greatly influence the orbit of spacecraft about the Moon. There are some puzzles: lava flows by themselves cannot explain all of the gravitational signature, and some mascons exist that are not linked to mare volcanism.

The Moon has an external magnetic field of the order of one to a hundred nanoteslas, less than one-hundredth that of the Earth. It does not currently have a global dipolar magnetic field, as would be generated by a liquid metal core geodynamo, and only has crustal magnetization, probably acquired early in lunar history when a geodynamo was still operating. Alternatively, some of the remnant magnetization may be from transient magnetic fields generated during large impact events, through the expansion of an impact-generated plasma cloud in the presence of an ambient magnetic field—this is supported by the apparent location of the largest crustal magnetizations near the antipodes of the giant impact basins.

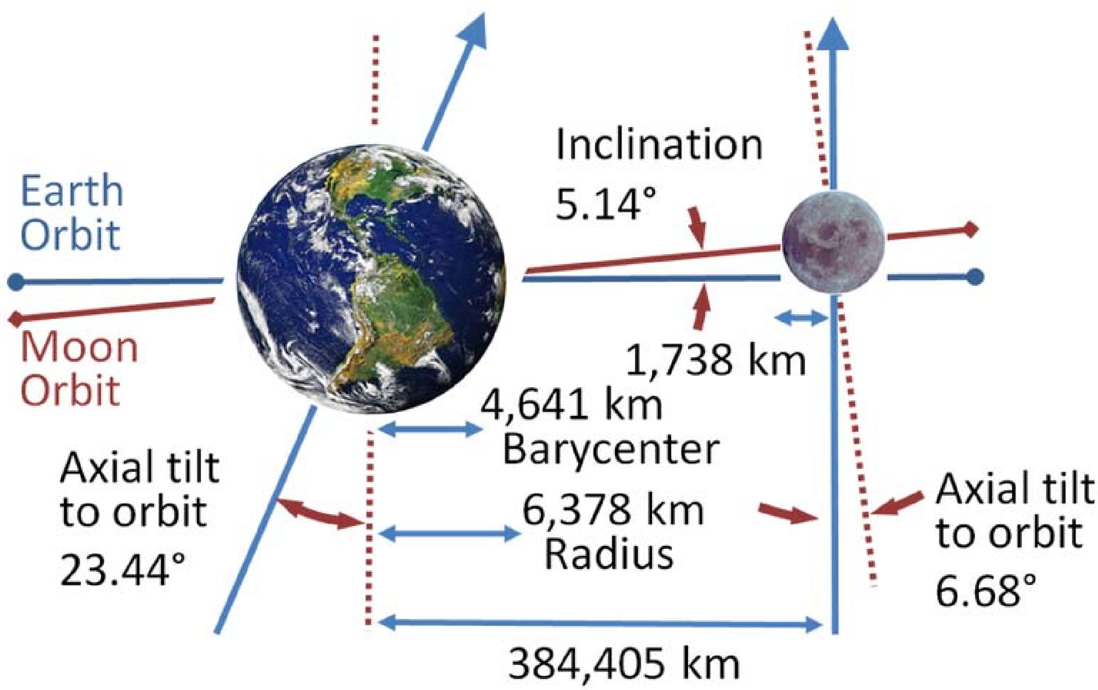
Atmosphere

The Moon has an atmosphere so tenuous as to be nearly vacuum, with a total mass of less than 10 metric tons. The surface pressure of this small mass is around 3×10^{-15} atm (0.3 nPa); it varies with the lunar day. Its sources include outgassing and sputtering, the release of atoms from the bombardment of lunar soil by solar wind ions. Elements that have been detected include sodium and potassium, produced by sputtering, which are also found in the atmospheres of Mercury and Io; helium-4 from the solar wind; and argon-40, radon-222, and polonium-210, outgassed after their creation by radioactive decay within the crust and mantle. The absence of such neutral species (atoms or molecules) as oxygen, nitrogen, carbon, hydrogen and magnesium, which are present in the regolith, is not understood. Water vapour has been detected by *Chandrayaan-1* and found to vary with latitude, with a maximum at ~60–70 degrees; it is possibly generated from the sublimation of water ice in the regolith. These gases can either return into the regolith due to the Moon's gravity, or be lost to space: either through solar radiation pressure, or if they are ionised, by being swept away by the solar wind's magnetic field.

Seasons

The Moon's axial tilt is only 1.54°, much less than the 23.44° of the Earth. Because of this, the Moon's solar illumination varies much less with season, and topographical details play a crucial role in seasonal effects. From images taken by *Clementine* in 1994, it appears that four mountainous regions on the rim of Peary crater at the Moon's north pole remain illuminated for the entire lunar day, creating peaks of eternal light. No such regions exist at the south pole. Similarly, there are places that remain in permanent shadow at the bottoms of many polar craters, and these dark craters are extremely cold: *Lunar Reconnaissance Orbiter* measured the lowest summer temperatures in craters at the southern pole at 35 K (−238 °C), and just 26 K close to the winter solstice in north polar Hermite Crater. This is the coldest temperature in the Solar System ever measured by a spacecraft, colder even than the surface of Pluto.

Relationship to Earth

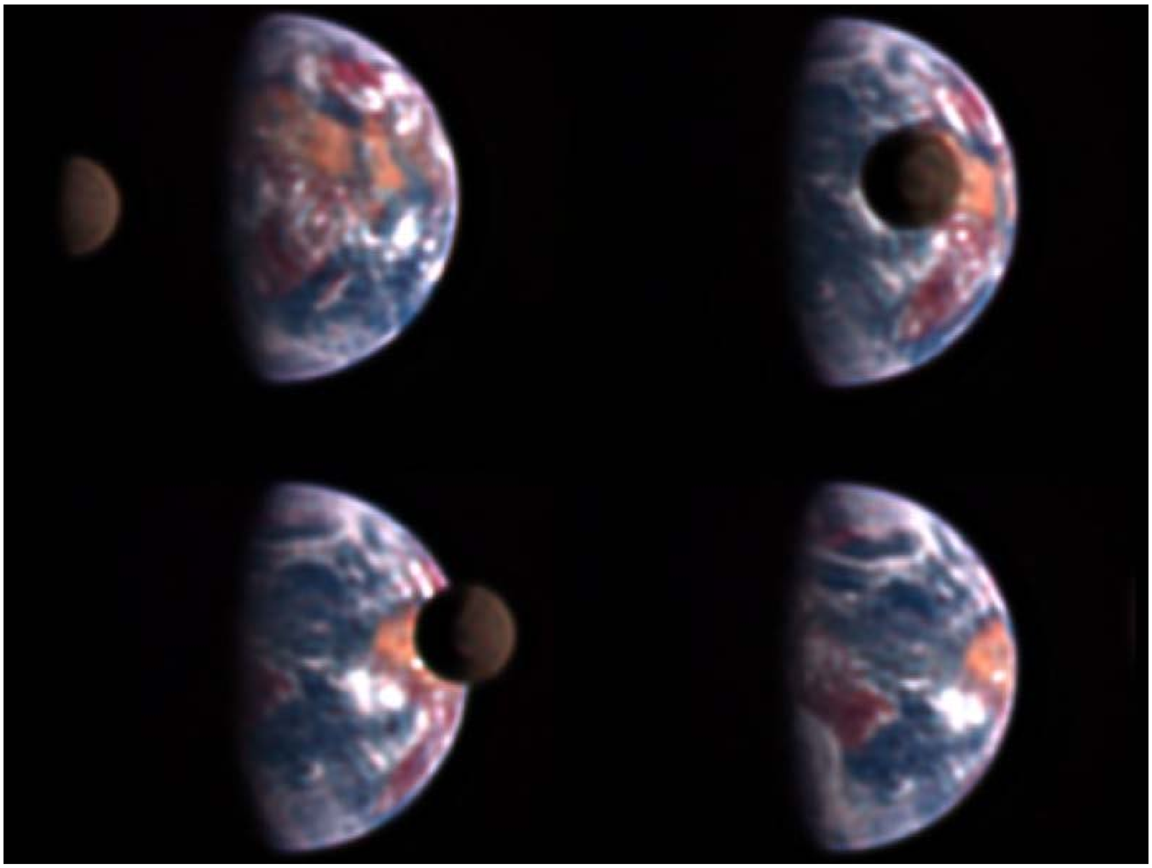


Schematic of the Earth-Moon system (without a consistent scale)

Orbit

The Moon makes a complete orbit around the Earth with respect to the fixed stars about once every 27.3 days (its sidereal period). However, since the Earth is moving in its orbit about the Sun at the same time, it takes slightly longer for the Moon to show the same phase to Earth, which is about 29.5 days (its synodic period). Unlike most satellites of other planets, the Moon orbits nearer the ecliptic plane than to the planet's equatorial plane. The Moon's orbit is subtly perturbed by the Sun and Earth in many small, complex and interacting ways. For example, the plane of the Moon's orbital motion gradually rotates, which affects other aspects of lunar motion. These follow-on effects are mathematically described by Cassini's laws.

Relative size



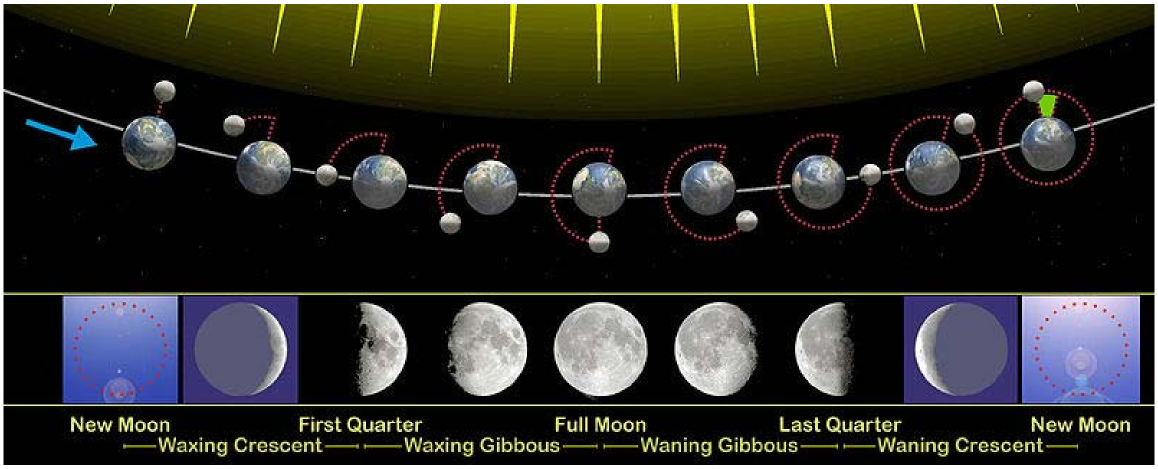
Comparative sizes of the Earth and the Moon, as imaged at separation of 50 million km

The Moon is exceptionally large relative to the Earth: a quarter the diameter of the planet and 1/81 its mass. It is the largest moon in the solar system relative to the size of its planet (although Charon is larger relative to the dwarf planet Pluto).

However, the Earth and Moon are still considered a planet–satellite system, rather than a double-planet system, as their barycentre, the common centre of mass, is located 1,700 km (about a quarter of the Earth's radius) beneath the surface of the Earth.

Appearance from Earth

The Moon has an exceptionally low albedo, giving it a similar reflectance to coal. Despite this, it is the second brightest object in the sky after the Sun. This is partly due to the brightness enhancement of the opposition effect; at quarter phase, the Moon is only one-tenth as bright, rather than half as bright, as at full Moon. Additionally, colour constancy in the visual system recalibrates the relations between the colours of an object and its surroundings, and since the surrounding sky is comparatively dark, the sunlit Moon is perceived as a bright object. The edges of the full Moon seem as bright as the centre, with no limb darkening, due to the reflective properties of lunar soil, which reflects more light back towards the Sun than in other directions. The Moon does appear larger when close to the horizon, but this is a purely psychological effect, known as the Moon illusion, first described in the 7th century BC.



The monthly changes of angle between the direction of illumination by the Sun and viewing from Earth, and the phases of the Moon that result

The highest altitude of the Moon in the sky varies: while it has nearly the same limit as the Sun, it alters with the lunar phase and with the season of the year, with the full Moon highest during winter. The 18.6-year nodes cycle also has an influence: when the ascending node of the lunar orbit is in the vernal equinox, the lunar declination can go as far as 28° each month. This means the Moon can go overhead at latitudes up to 28° from the equator, instead of only 18°. The orientation of the Moon's crescent also depends on the latitude of the observation site: close to the equator, an observer can see a smile-shaped crescent Moon.

There has been historical controversy over whether features on the Moon's surface change over time. Today, many of these claims are thought to be illusory, resulting from observation under different lighting conditions, poor astronomical seeing, or inadequate drawings. However, outgassing does occasionally occur, and could be responsible for a minor percentage of the reported lunar transient phenomena. Recently, it has been suggested that a roughly 3 km diameter region of the lunar surface was modified by a gas release event about a million years ago. The Moon's appearance, like that of the Sun, can be affected by Earth's atmosphere: common effects are a 22° halo ring formed when the Moon's light is refracted through the ice crystals of high cirrostratus cloud, and smaller coronal rings when the Moon is seen through thin clouds.

Tidal effects

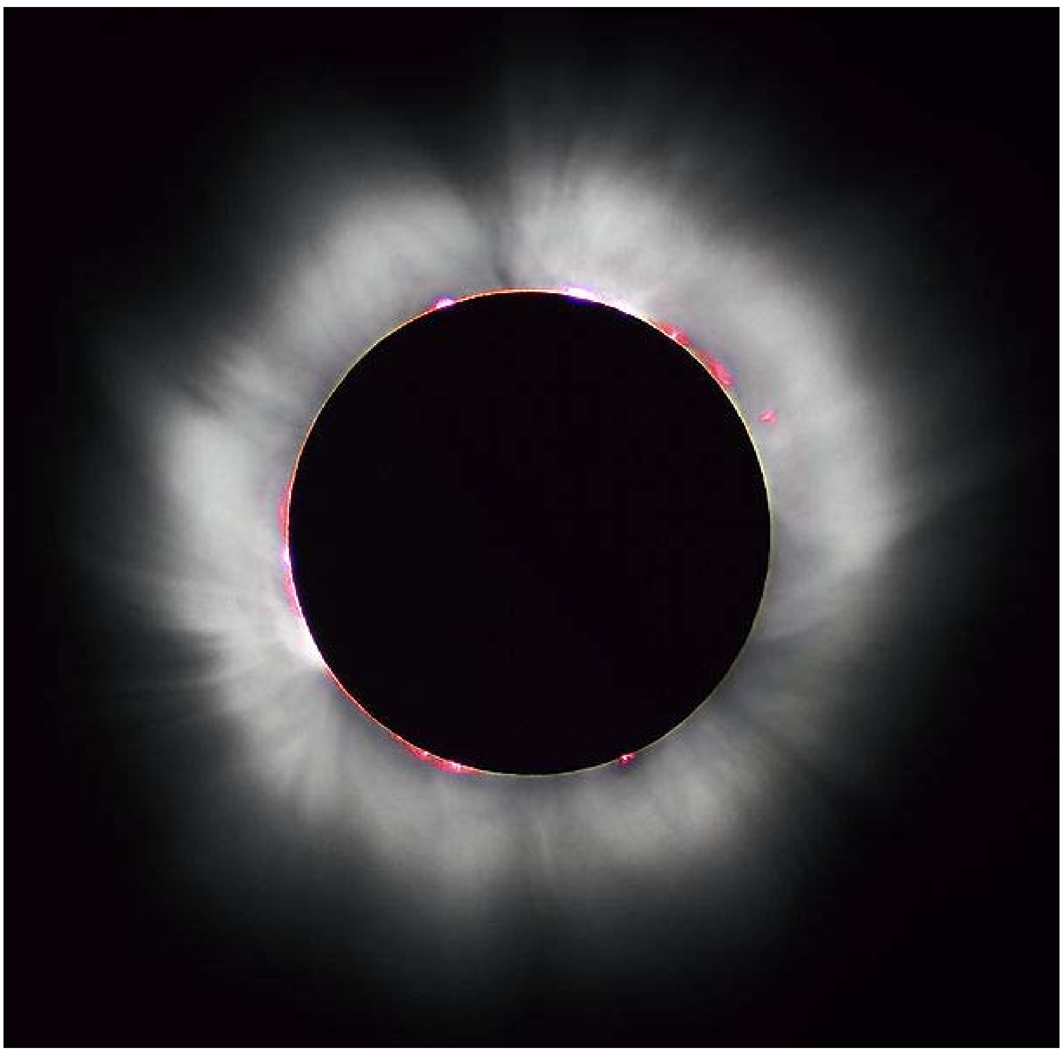
The tides on the Earth are mostly generated by the gradient in intensity of the Moon's gravitational pull from one side of the Earth to the other, the tidal forces. This forms two tidal bulges on the Earth, which are most clearly seen in elevated sea level as ocean tides. Since the Earth spins about 27 times faster than the Moon moves around it, the bulges are dragged along with the Earth's surface faster than the Moon moves, rotating around the Earth once a day as it spins on its axis. The ocean tides are magnified by other effects: frictional coupling of water to Earth's rotation through the ocean floors, the inertia of water's movement, ocean basins that get shallower near land, and oscillations between different ocean basins. The gravitational attraction of the Sun on the Earth's oceans is almost half that of the Moon, and their gravitational interplay is responsible for spring and neap tides.

The libration of the Moon over a single lunar month

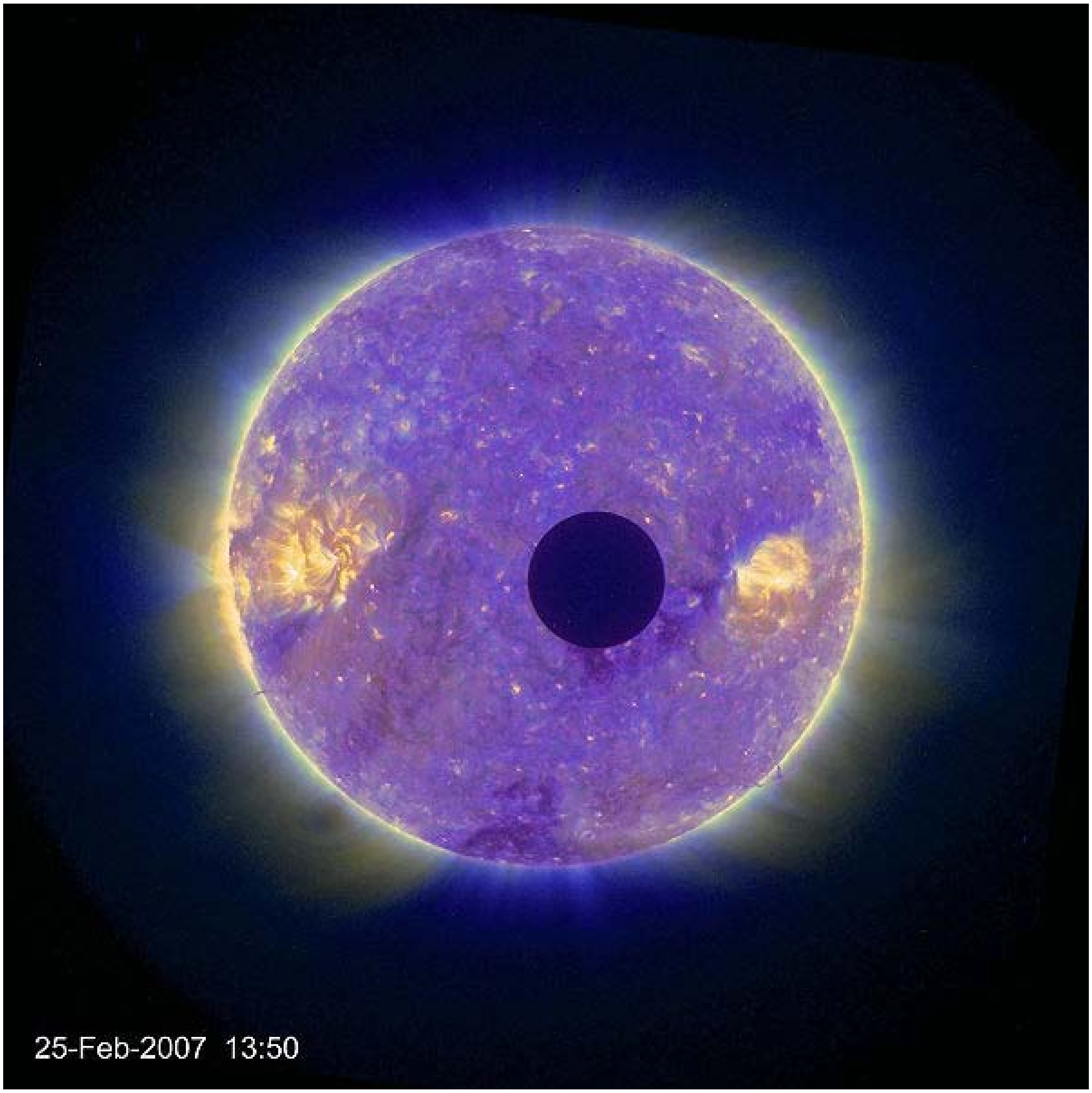
Gravitational coupling between the Moon and the bulge nearest the Moon acts as a torque on the Earth's rotation, draining angular momentum and rotational kinetic energy from the Earth's spin. In turn, angular momentum is added to the Moon's orbit, accelerating it, which lifts the Moon into a higher orbit with a longer period. As a result, the distance between the Earth and Moon is increasing, and the Earth's spin slowing down. Measurements from lunar ranging experiments with laser reflectors left during the Apollo missions have found that the Moon's distance to the Earth increases by 38 mm per year (though this is only 0.10 ppb/year of the radius of the Moon's orbit). Atomic clocks also show that the Earth's day lengthens by about 15 microseconds every year, slowly increasing the rate at which UTC is adjusted by leap seconds. This tidal drag will continue until the spin of the Earth has slowed to match the orbital period of the Moon; however, long before this could happen, the Sun will have become a red giant, engulfing the Earth.

The lunar surface also experiences tides of amplitude ~10 cm over 27 days, with two components: a fixed one due to the Earth, as they are in synchronous rotation, and a varying component from the Sun. The Earth-induced component arises from libration, a result of the Moon's orbital eccentricity; if the Moon's orbit were perfectly circular, there would only be solar tides. Libration also changes the angle from which the Moon is seen, allowing about 59% of its surface to be seen from the Earth (but only half at any instant). The cumulative effects of stress built up by these tidal forces produces *moonquakes*. Moonquakes are much less common and weaker than earthquakes, although they can last for up to an hour - a significantly longer time than terrestrial earthquakes - because of the absence of water to damp out the seismic vibrations. The existence of moonquakes was an unexpected discovery from seismometers placed on the Moon by Apollo astronauts from 1969 through 1972.

Eclipses



The 1999 solar eclipse



The Moon passing in front of the Sun, from the STEREO-B spacecraft

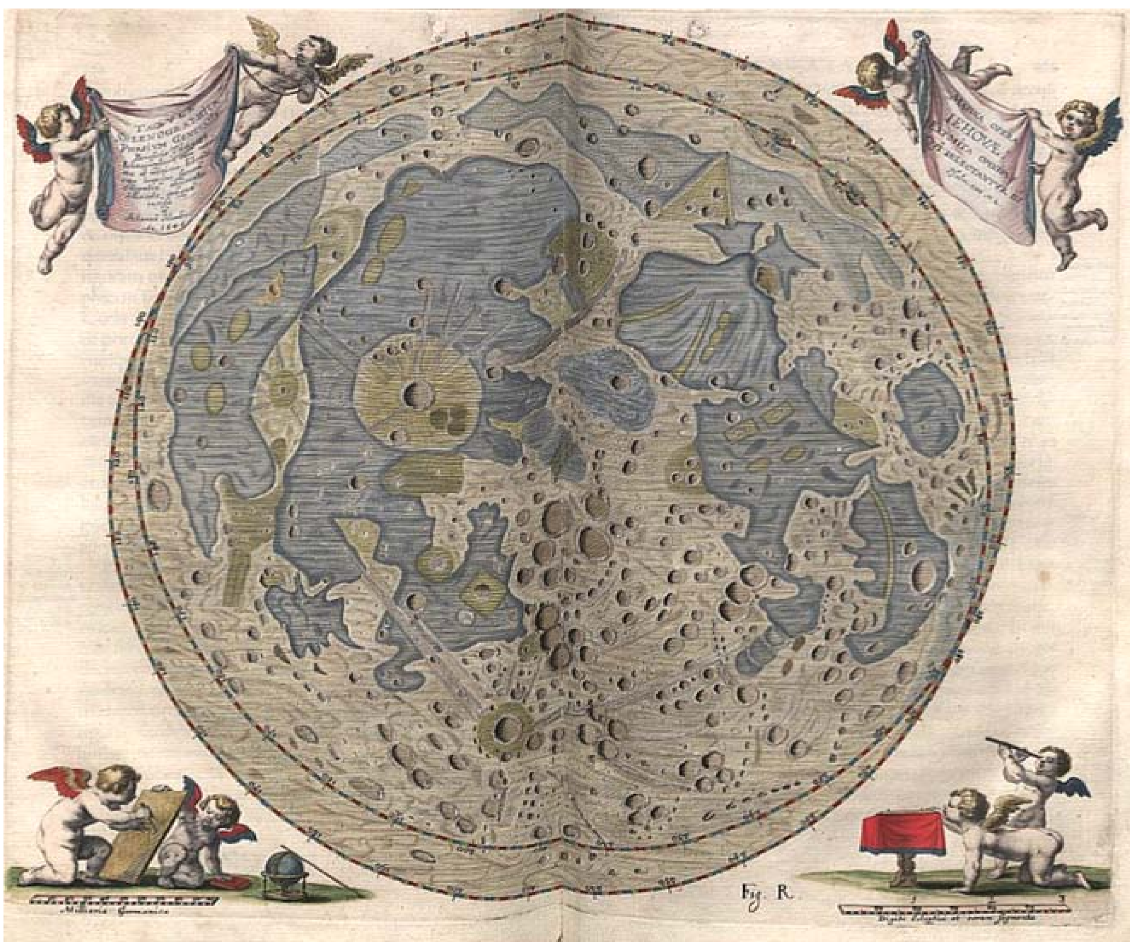
From the Earth, the Moon and Sun appear the same size. From a satellite in an Earth-trailing orbit, the Moon appears smaller than the Sun.

Eclipses can only occur when the Sun, Earth, and Moon are all in a straight line. Solar eclipses occur near a new Moon, when the Moon is between the Sun and Earth. In contrast, lunar eclipses occur near a full Moon, when the Earth is between the Sun and Moon. The angular diameters of the Moon and the Sun as seen from Earth overlap in their variation, so that both total and annular solar eclipses are possible. In a total eclipse, the Moon completely covers the disc of the Sun and the solar corona becomes visible to the naked eye. Since the distance between the Moon and the Earth is very slowly increasing over time, the angular diameter of the Moon is decreasing. This means that hundreds of millions of years ago the Moon would always completely cover the Sun on solar eclipses, and no annular eclipses were possible. Likewise, about 600 million years from now (if the angular diameter of the Sun does not change), the Moon will no longer cover the Sun completely, and only annular eclipses will occur.

Because the Moon's orbit around the Earth is inclined by about 5° to the orbit of the Earth around the Sun, eclipses do not occur at every full and new Moon. For an eclipse to occur, the Moon must be near the intersection of the two orbital planes. The periodicity and recurrence of eclipses of the Sun by the Moon, and of the Moon by the Earth, is described by the saros cycle, which has a period of approximately 18 years.

As the Moon is continuously blocking our view of a half-degree-wide circular area of the sky, the related phenomenon of occultation occurs when a bright star or planet passes behind the Moon and is occulted: hidden from view. In this way, a solar eclipse is an occultation of the Sun. Because the Moon is comparatively close to the Earth, occultations of individual stars are not visible everywhere on the planet, nor at the same time. Because of the precession of the lunar orbit, each year different stars are occulted.

Study and exploration



Map of the Moon by Johannes Hevelius from his Selenographia (1647), the first map to include the libration zones.

Early studies

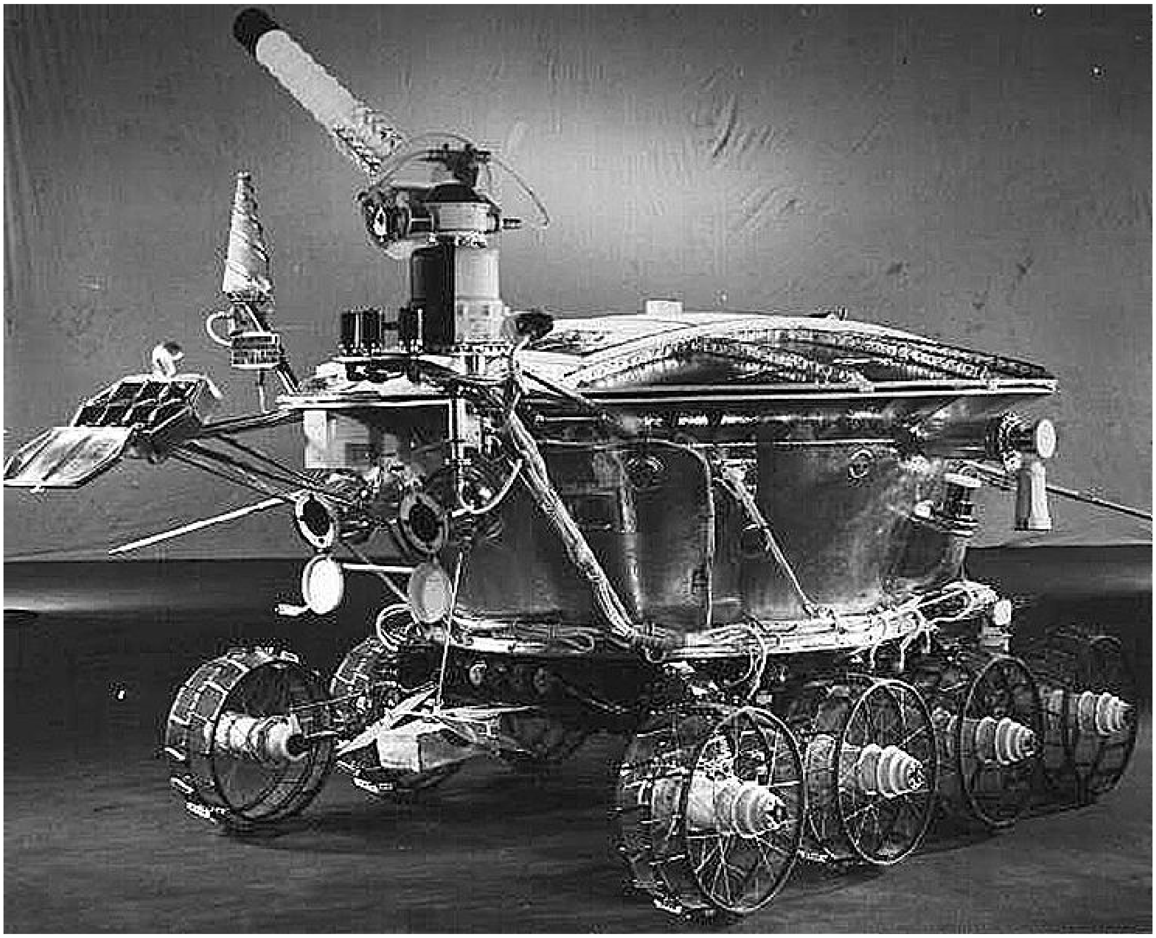
Understanding of the Moon's cycles was an early development of astronomy: by the 5th century BC, Babylonian astronomers had recorded the 18-year Saros cycle of lunar eclipses, and Indian astronomers had described the Moon's monthly elongation. The Chinese astronomer Shi Shen (fl. 4th century BC) gave instructions for predicting solar and lunar eclipses. Later, the physical form of the Moon and the cause of moonlight became understood. The ancient Greek philosopher Anaxagoras (d. 428 BC) reasoned that the Sun and Moon were both giant spherical rocks, and that the latter reflected the light of the former. Although the Chinese of the Han Dynasty believed the Moon to be energy equated to *qi*, their 'radiating influence' theory also recognized that the light of the Moon was merely a reflection of the Sun, and Jing Fang (78–37 BC) noted the sphericity of the Moon. In 499 AD, the Indian astronomer Aryabhata mentioned in his *Aryabhatiya* that reflected sunlight is the cause of the shining of the Moon. Shen Kuo (1031–1095) of the Song Dynasty created an allegory equating the waxing and waning of the Moon to a round ball of reflective silver that, when doused with white powder and viewed from the side, would appear to be a crescent.

In Aristotle's (384–322 BC) description of the universe, the Moon marked the boundary between the spheres of the mutable elements (earth, water, air and fire), and the imperishable stars of aether, an influential philosophy that would dominate for centuries. However, in the 2nd century BC, Seleucus of Seleucia correctly theorized that tides were due to the attraction of the Moon, and that their height depends on the Moon's position relative to the Sun. In the same century, Aristarchus computed the size and distance of the Moon from Earth, obtaining a value of about twenty times the Earth radius for the distance. These figures were greatly improved by Ptolemy (90–168 AD): his values of a mean distance of 59 times the Earth's radius and a diameter of 0.292 Earth diameters were close to the correct values of about 60 and 0.273 respectively. Archimedes (287–212 BC) invented a planetarium calculating motions of the Moon and the known planets.

During the Middle Ages, before the invention of the telescope, the Moon was increasingly recognised as a sphere, though many believed that it was "perfectly smooth". In 1609, Galileo Galilei drew one of the first telescopic drawings of the Moon in his book *Sidereus Nuncius* and noted that it was not smooth but had mountains and craters. Telescopic mapping of the Moon followed: later in the 17th century, the efforts of Giovanni Battista Riccioli and Francesco Maria Grimaldi led to the system of naming of lunar features in use today. The more exact 1834-6 *Mappa Selenographica* of Wilhelm Beer and Johann Heinrich Mädler, and their associated 1837 book *Der Mond*, the first trigonometrically accurate study of lunar features, included the heights of more than a thousand mountains, and introduced the study of the Moon at accuracies possible in earthly geography. Lunar craters, first noted by Galileo, were thought to be volcanic until the 1870s proposal of Richard Proctor that they were formed by collisions. This view gained support in 1892 from the experimentation of geologist Grove Karl Gilbert, and from comparative studies from 1920 to the 1940s, leading to the development of lunar stratigraphy, which by the 1950s was becoming a new and growing branch of astrogeology.

First direct exploration: 1959–1976

Soviet missions



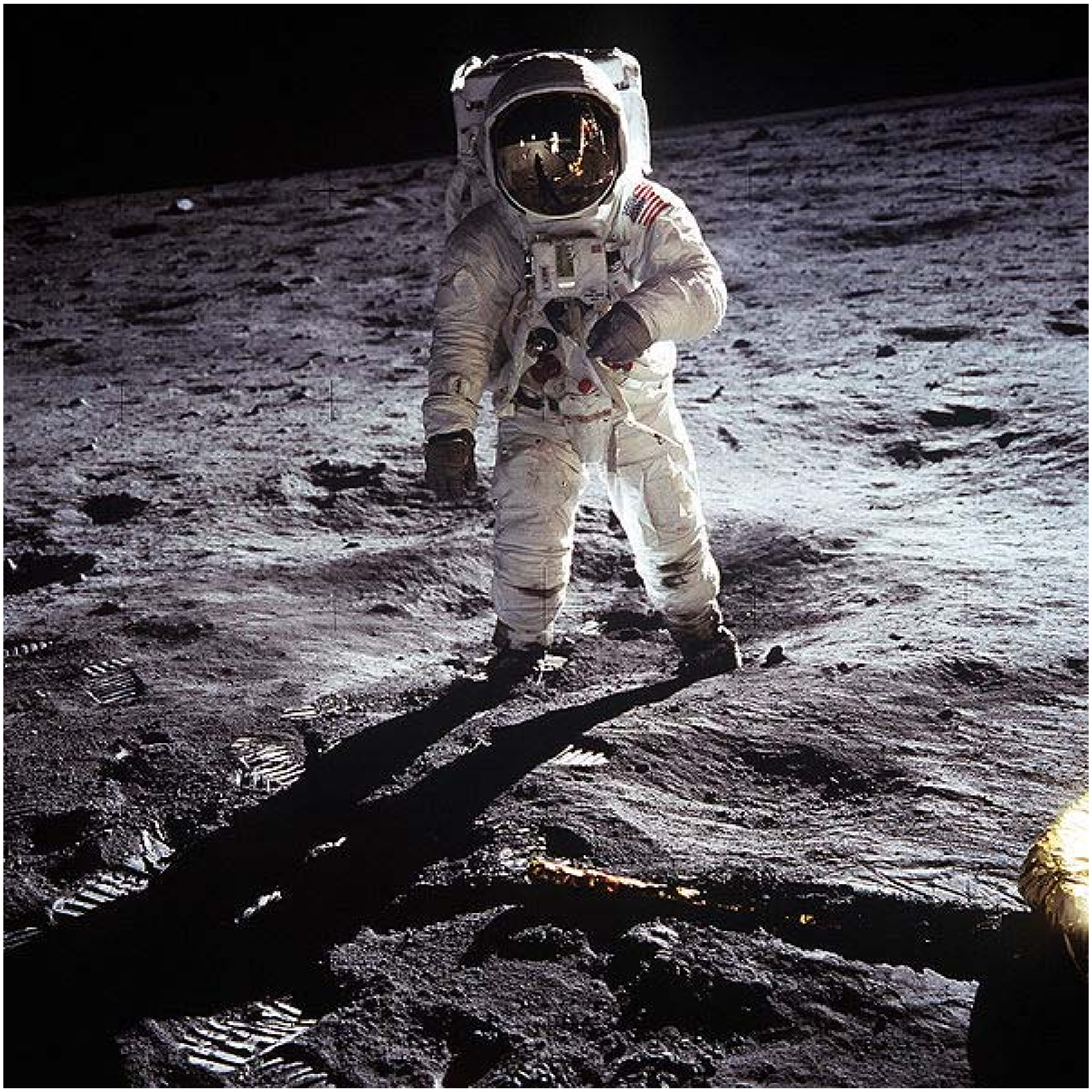
Lunokhod 1 (lit. *moonwalker*), the first successful space rover

The Cold War-inspired space race between the Soviet Union and the U.S. led to an acceleration of interest in exploration of the Moon. Once launchers had the necessary capabilities, these nations sent unmanned probes on both flyby and impact/lander missions. Spacecraft from the Soviet Union's *Luna* program were the first to accomplish a number of goals: following three unnamed, failed missions in 1958, the first man-made object to escape Earth's gravity and pass near the Moon was *Luna 1*; the first man-made object to impact the lunar surface was *Luna 2*, and the first photographs of the normally occluded far side of the Moon were made by *Luna 3*, all in 1959.

The first spacecraft to perform a successful lunar soft landing was *Luna 9* and the first unmanned vehicle to orbit the Moon was *Luna 10*, both in 1966. Rock and soil samples were brought back to Earth by three *Luna* sample return missions (*Luna 16*, *20*, and *24*), which returned 0.3 kg total. Two pioneering robotic spacecrafts of rover type landed on the Moon in 1970 and 1973 as a part of Soviet Lunokhod programme.

United States missions

Earth as viewed from the Moon during the Apollo 8 mission, Christmas Eve, 1968. Africa is at the sunset terminator, both Americas are under cloud, and Antarctica is at the left end of the terminator.



Astronaut Buzz Aldrin photographed by Neil Armstrong during the first Moon landing on 20 July 1969

American lunar exploration began with robotic missions aimed at developing understanding of the lunar surface for an eventual manned landing: the Jet Propulsion Laboratory's *Surveyor* program landed its first spacecraft four months after *Luna 9*. NASA's manned Apollo program was developed in parallel; after a series of unmanned and manned tests of the Apollo spacecraft in Earth orbit, and spurred on by a potential Soviet lunar flight, in 1968 Apollo 8 made the first crewed mission to lunar orbit. The subsequent landing of the first humans on the Moon in 1969 is seen by many as the culmination of the space race. Neil Armstrong became the first person to walk on the Moon as the commander of the American mission Apollo 11 by first setting foot on the Moon at 02:56 UTC on 21 July 1969. The Apollo missions 11 to 17 (except Apollo 13, which aborted its planned lunar landing) returned 382 kg of lunar rock and soil in 2,196 separate samples. The American Moon landing and return was enabled by considerable technological advances in the early 1960s, in domains such as ablation chemistry, software engineering and atmospheric re-entry technology, and by highly competent management of the enormous technical undertaking.

Scientific instrument packages were installed on the lunar surface during all the Apollo missions. Long-lived instrument stations, including heat flow probes, seismometers, and magnetometers, were installed at the Apollo 12, 14, 15, 16, and 17 landing sites. Direct transmission of data to Earth concluded in late 1977 due to budgetary considerations, but as the stations' lunar laser ranging corner-cube retroreflector arrays are passive instruments, they are still being used. Ranging to the stations is routinely performed from earth-based stations with an accuracy of a few centimetres, and data from this experiment are being used to place constraints on the size of the lunar core.

Current era: 1990–present

Post-Apollo and *Luna*, many more countries have become involved in direct exploration of the Moon. In 1990, Japan orbited the Moon with the *Hiten* spacecraft, becoming the third country to place a spacecraft into lunar orbit. The spacecraft released a smaller probe, *Hagoromo*, in lunar orbit, but the transmitter failed, preventing further scientific use of the mission. In 1994, the U.S. sent the joint Defense Department/NASA spacecraft *Clementine* to lunar orbit. This mission obtained the first near-global topographic map of the Moon, and the first global multispectral images of the lunar surface. This was followed in 1998 by the *Lunar Prospector* mission, whose instruments indicated the presence of excess hydrogen at the lunar poles, which is likely to have been caused by the presence of water ice in the upper few meters of the regolith within permanently shadowed craters.

The European spacecraft *Smart 1*, the second ion-propelled spacecraft, was in lunar orbit from 15 November 2004 until its lunar impact on 3 September 2006, and made the first detailed survey of chemical elements on the lunar surface. China has expressed ambitious plans for exploring the Moon, and successfully orbited its first spacecraft, *Chang'e-1*, from 5 November 2007 until its controlled lunar impact on 1 March 2008. In its sixteen-month mission, it obtained a full image map of the Moon. Between 4 October 2007 and 10 June 2009, the Japan Aerospace Exploration Agency's *Kaguya (Selene)* mission, a lunar orbiter fitted with a high-definition video camera, and two small radio-transmitter satellites, obtained lunar geophysics data and took the first high-definition movies from beyond Earth orbit. India's first lunar mission, *Chandrayaan I*, orbited from 8 November 2008 until loss of contact on 27 August 2009, creating a high resolution chemical, mineralogical and photo-geological map of the lunar surface, and confirming the presence of water molecules in lunar soil. The Indian Space Research Organisation plans to launch *Chandrayaan II* in 2013, which is slated to include a Russian robotic lunar rover. The U.S. co-launched the *Lunar Reconnaissance Orbiter* (LRO) and the *LCROSS* impactor and follow-up observation orbiter on 18 June 2009; *LCROSS* completed its mission by making a planned and widely observed impact in the crater Cabeus on 9 October 2009, while *LRO* is currently in operation, obtaining precise lunar altimetry and high-resolution imagery.

Other upcoming lunar missions include Russia's *Luna-Glob*: an unmanned lander, set of seismometers, and an orbiter based on its Martian *Phobos-Grunt* mission, which is slated to launch in 2012. Privately funded lunar exploration has been promoted by the Google Lunar X Prize, announced 13 September 2007, which offers US$20 million to anyone who can land a robotic rover on the Moon and meet other specified criteria.

NASA began to plan to resume manned missions following the call by U.S. President George W. Bush on 14 January 2004 for a mission to the Moon by 2020. The Constellation program was funded and construction and testing begun on a manned spacecraft and launch vehicle, and design studies for a lunar base. However, that program has been placed in jeopardy by the proposed 2011 budget, which will cancel Constellation in favour of NASA pursuing space technology and heavy-lift rocketry research. India has also expressed its hope for a manned mission to the Moon by 2020.

Legal status

Although *Luna* landers scattered pennants of the Soviet Union on the Moon, and U.S. flags were symbolically planted at their landing sites by the Apollo astronauts, no nation currently claims ownership of any part of the Moon's surface. Russia and the U.S. are party to the 1967 Outer Space Treaty, which defines the Moon and all outer space as the "province of all mankind". This treaty also restricts the use of the Moon to peaceful purposes, explicitly banning military installations and weapons of mass destruction. The 1979 Moon Agreement was created to restrict the exploitation of the Moon's resources by any single nation, but it has not been signed by any of the space-faring nations. While several individuals have made claims to the Moon in whole or in part, none of these are considered credible.

Chapter- 5

Abiogenesis

Pre-Cambrian stromatolites in the Siyeh Formation, Glacier National Park. In 2002, William Schopf of UCLA published a paper in the scientific journal *Nature* arguing that geological formations such as this possess 3.5 Ga (billion years old) fossilized cyanobacteria microbes. If true, they would be evidence of the earliest known life on earth.

In natural science, **abiogenesis** or **biopoesis** is the study of how life on Earth could have arisen from inanimate matter. Most amino acids, often called "the building blocks of life", can form via natural chemical reactions unrelated to life, as demonstrated in the Miller–Urey experiment and similar experiments, which involved simulating some of the conditions of the early Earth, in a scientific laboratory. In all living things, these amino acids are organized into proteins, and the construction of these proteins is mediated by nucleic acids. Which of these organic molecules first arose and how they formed the first life is the focus of abiogenesis.

In any theory of abiogenesis, two aspects of life have to be accounted for: replication, and metabolism. The question of which came first gave rise to different types of theories. In the beginning, metabolism-first theories (Oparin coacervate) were proposed, and only later thinking gave rise to modern, replication-first approach.

In modern, still somewhat limited understanding, the first living things on Earth are thought to be single cell prokaryotes (which lack a cell nucleus), perhaps evolved from protobionts (organic molecules surrounded by a membrane-like structure). The oldest ancient fossil microbe-like objects are dated to be 3.5 Ga (billion years old), approximately one billion years after the formation of the Earth itself. By 2.4 Ga, the ratio of stable isotopes of carbon, iron and sulfur shows the action of living things on inorganic minerals and sediments and molecular biomarkers indicate photosynthesis, demonstrating that life on Earth was widespread by this time.

The sequence of chemical events that led to the first nucleic acids is not known. Several hypotheses about early life have been proposed, most notably the iron-sulfur world theory (metabolism without genetics) and the RNA world hypothesis (RNA life-forms).

Conceptual history

Spontaneous generation

Until the early 19th century, people generally believed in the ongoing spontaneous generation of certain forms of life from non-living matter. This was paired with heterogenesis, the belief that one form of life derives from a different form (*e.g.* bees from flowers). Classical notions of abiogenesis, now more precisely known as *spontaneous generation*, held that certain complex, living organisms are generated by decaying organic substances. According to Aristotle it was a readily observable truth that aphids arise from the dew which falls on plants, fleas from putrid matter, mice from dirty hay, crocodiles from rotting logs at the bottom of bodies of water, and so on.

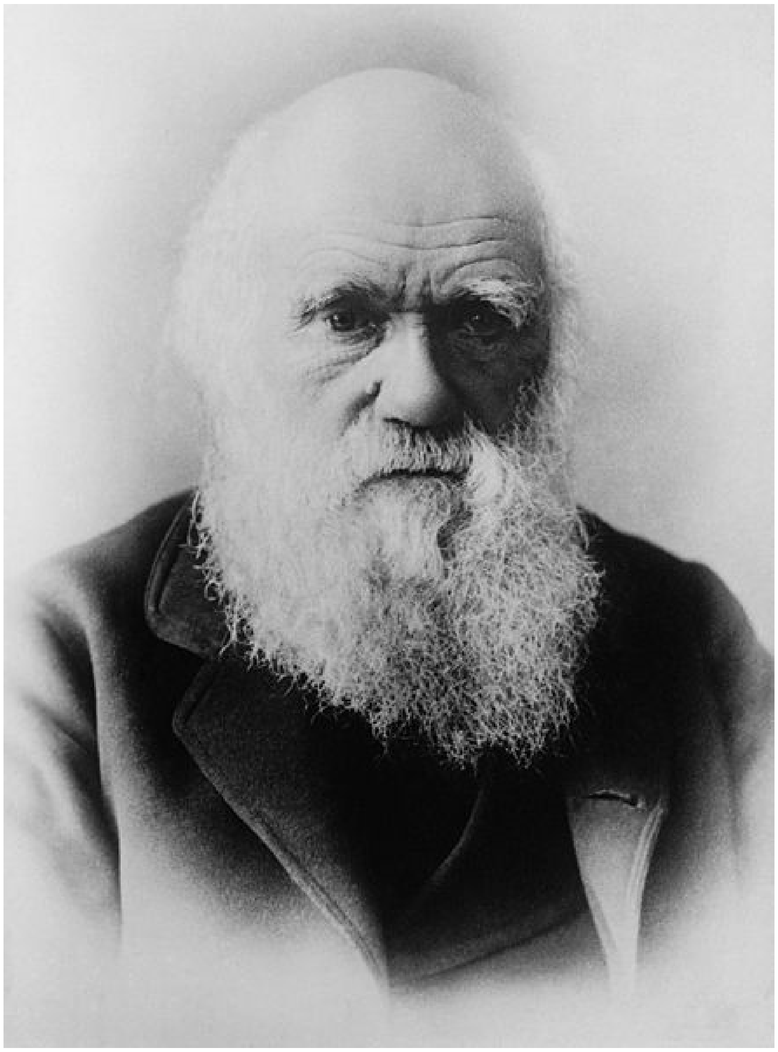
In the 17th century, such assumptions started to be questioned; for example, in 1646, Sir Thomas Browne published his *Pseudodoxia Epidemica* (subtitled *Enquiries into Very many Received Tenets, and Commonly Presumed Truths*), which was an attack on false beliefs and "vulgar errors." His conclusions were not widely accepted. For example, his contemporary, Alexander Ross wrote: "To question this (i.e., spontaneous generation) is to question reason, sense and experience. If he doubts of this let him go to Egypt, and there he will find the fields swarming with mice, begot of the mud of Nylus, to the great calamity of the inhabitants."

In 1665, Robert Hooke published the first drawings of a microorganism. Hooke was followed in 1676 by Anton van Leeuwenhoek, who drew and described microorganisms that are now thought to have been protozoa and bacteria. Many felt the existence of microorganisms was evidence in support of spontaneous generation, since microorganisms seemed too simplistic for sexual reproduction, and asexual reproduction through cell division had not yet been observed.

The first solid evidence against spontaneous generation came in 1668 from Francesco Redi, who proved that no maggots appeared in meat when flies were prevented from laying eggs. It was gradually shown that, at least in the case of all the higher and readily visible organisms, the previous sentiment regarding spontaneous generation was false. The alternative seemed to be biogenesis: that every living thing came from a pre-existing living thing (*omne vivum ex ovo*, Latin for "every living thing from an egg").

In 1768, Lazzaro Spallanzani demonstrated that microbes were present in the air, and could be killed by boiling. In 1861, Louis Pasteur performed a series of experiments which demonstrated that organisms such as bacteria and fungi do not spontaneously appear in sterile, nutrient-rich media.

Pasteur and Darwin



Charles Darwin in 1879

By the middle of the 19th century, the theory of biogenesis had accumulated so much evidential support, due to the work of Louis Pasteur and others, that the alternative theory of spontaneous generation had been effectively disproven. Pasteur himself remarked, after a definitive finding in 1864, "Never will the doctrine of spontaneous generation recover from the mortal blow struck by this simple experiment." The collapse of spontaneous generation, however, left a vacuum of scientific thought on the question of how life *had* first arisen.

In a letter to Joseph Dalton Hooker on February 1, 1871, Charles Darwin addressed the question, suggesting that the original spark of life may have begun in a "warm little pond, with all sorts of ammonia and phosphoric salts, lights, heat, electricity, etc. present, so that a protein compound was chemically formed ready to undergo still more complex changes". He went on to explain that "at the present day such matter would be instantly devoured or absorbed, which would not have been the case before living creatures were formed." In other words, the presence of life itself makes the search for the origin of life dependent on the sterile conditions of the laboratory.

"Primordial soup" theory

Alexander Oparin (right) at the laboratory

No new notable research or theory on the subject appeared until 1924, when Alexander Oparin reasoned that atmospheric oxygen prevents the synthesis of certain organic compounds that are necessary building blocks for the evolution of life. In his *The Origin of Life*, Oparin proposed that the "spontaneous generation of life" that had been attacked by Louis Pasteur, did in fact occur once, but was now impossible because the conditions found in the early earth had changed, and the presence of living organisms would immediately consume any spontaneously generated organism. Oparin argued that a "primeval soup" of organic molecules could be created in an oxygen-less atmosphere through the action of sunlight. These would combine in ever-more complex fashions until they formed coacervate droplets. These droplets would "grow" by fusion with other droplets, and "reproduce" through fission into daughter droplets, and so have a primitive metabolism in which those factors which promote "cell integrity" survive, and those that do not become extinct. Many modern theories of the origin of life still take Oparin's ideas as a starting point.

Around the same time, J. B. S. Haldane suggested that the Earth's pre-biotic oceans–very different from their modern counterparts–would have formed a "hot dilute soup" in which organic compounds could have formed. This idea was called *biopoiesis* or *biopoesis*, the process of living matter evolving from self-replicating but nonliving molecules.

Early conditions

Morse and MacKenzie have suggested that oceans may have appeared first in the Hadean eon, as soon as two hundred million years (200 Ma) after the Earth was formed, in a hot 100 °C (212 °F) reducing environment, and that the pH of about 5.8 rose rapidly towards neutral. This has been supported by Wilde who has pushed the date of the zircon crystals found in the metamorphosed quartzite of Mount Narryer in Western Australia, previously thought to be 4.1–4.2 Ga, to 4.404 Ga. This means that oceans and continental crust existed within 150 Ma of Earth's formation.

Despite this, the Hadean environment was one highly hazardous to life. Frequent collisions with large objects, up to 500 kilometres (310 mi) in diameter, would have been sufficient to vaporise the ocean within a few months of impact, with hot steam mixed with rock vapour leading to high altitude clouds completely covering the planet. After a few months the height of these clouds would have begun to decrease but the cloud base would still have been elevated for about the next thousand years. After that, it would have begun to rain at low altitude. For another two thousand years rains would slowly have drawn down the height of the clouds, returning the oceans to their original depth only 3,000 years after the impact event.

Between 3.8 and 4.1 Ga, changes in the orbits of the gaseous giant planets may have caused a late heavy bombardment that pockmarked the moon and other inner planets (Mercury, Mars, and presumably Earth and Venus). This would likely have sterilized the planet had life appeared before that time.

By examining the time interval between such devastating environmental events, the time interval when life might first have come into existence can be found for different early environments. The study by Maher and Stevenson shows that if the deep marine hydrothermal setting provides a suitable site for the origin of life, abiogenesis could have happened as early as 4.0 to 4.2 Ga, whereas if it occurred at the surface of the earth abiogenesis could only have occurred between 3.7 and 4.0 Ga.

Other research suggests a colder start to life. Work by Leslie Orgel and colleagues on the synthesis of purines has shown that freezing temperatures are advantageous, due to the concentrating effect for key precursors such as HCN. Research by Stanley Miller and colleagues suggested that while adenine and guanine require freezing conditions for synthesis, cytosine and uracil may require boiling temperatures. Based on this research, Miller suggested a beginning of life involving freezing conditions and exploding meteorites. An article in Discover Magazine points to research by the Miller group indicating the formation of seven different amino acids and 11 types of nucleobases in ice when ammonia and cyanide were left in a freezer from 1972–1997.

Here we, also describes research by Christof Biebricher showing the formation of RNA molecules 400 bases long under freezing conditions using an RNA template, a single-strand chain of RNA that guides the formation of a new strand of RNA. As that new RNA strand grows, it adheres to the template. The explanation given for the unusual speed of these reactions at such a low temperature is eutectic freezing. As an ice crystal forms, it stays pure: only molecules of water join the growing crystal, while impurities like salt or cyanide are excluded. These impurities become crowded in microscopic pockets of liquid within the ice, and this crowding causes the molecules to collide more often.

Evidence of the early appearance of life comes from the Isua supercrustal belt in Western Greenland and from similar formations in the nearby Akilia Islands. Carbon entering into rock formations has a ratio of Carbon-13 (^{13}C) to Carbon-12 (^{12}C) of about −5.5 (in units of $\delta^{13}C$), where because of a preferential biotic uptake of ^{12}C, biomass has a $\delta^{13}C$ of between −20 and −30. These isotopic fingerprints are preserved in the sediments, and Mojzis has used this technique to suggest that life existed on the planet already by 3.85 billion years ago. Lazcano and Miller (1994) suggest that the rapidity of the evolution of life is dictated by the rate of recirculating water through mid-ocean submarine vents. Complete recirculation takes 10 million years, thus any organic compounds produced by then would be altered or destroyed by temperatures exceeding 300 °C (572 °F). They estimate that the development of a 100 kilobase genome of a DNA/protein primitive heterotroph into a 7000 gene filamentous cyanobacterium would have required only 7 Ma.

Current models

There is no truly "standard model" of the origin of life. Most currently accepted models draw at least some elements from the framework laid out by the Oparin-Haldane hypothesis. Under that umbrella, however, are a wide array of disparate discoveries and conjectures such as the following, listed in a rough order of postulated emergence:

1. Some theorists suggest that the atmosphere of the early Earth may have been chemically reducing in nature, composed primarily of methane (CH_4), ammonia (NH_3), water (H_2O), hydrogen sulfide (H_2S), carbon dioxide (CO_2) or carbon monoxide (CO), and phosphate (PO_4^{3-}), with molecular oxygen (O_2) and ozone (O_3) either rare or absent.
2. In such a reducing atmosphere, electrical activity can catalyze the creation of certain basic small molecules (monomers) of life, such as amino acids. This was demonstrated in the Miller–Urey experiment by Stanley L. Miller and Harold C. Urey in 1953.
3. Phospholipids (of an appropriate length) can spontaneously form lipid bilayers, a basic component of the cell membrane.
4. A fundamental question is about the nature of the first self-replicating molecule. Since replication is accomplished in modern cells through the cooperative action of proteins and nucleic acids, the major schools of thought about how the process originated can be broadly classified as "proteins first" and "nucleic acids first".
5. The principal thrust of the "nucleic acids first" argument is as follows:
    1. The polymerization of nucleotides into random RNA molecules might have resulted in self-replicating ribozymes (RNA world hypothesis)
    2. Selection pressures for catalytic efficiency and diversity might have resulted in ribozymes which catalyse peptidyl transfer (hence formation of small proteins), since oligopeptides complex with RNA to form better catalysts. The first

ribosome might have been created by such a process, resulting in more prevalent protein synthesis.

3. Synthesized proteins might then outcompete ribozymes in catalytic ability, and therefore become the dominant biopolymer, relegating nucleic acids to their modern use, predominantly as a carrier of genomic information.

As of 2010, no one has yet synthesized a "protocell" using basic components which would have the necessary properties of life (the so-called *"bottom-up-approach"*). Without such a proof-of-principle, explanations have tended to be short on specifics. However, some researchers are working in this field, notably Steen Rasmussen at Los Alamos National Laboratory and Jack Szostak at Harvard University. Others have argued that a *"top-down approach"* is more feasible. One such approach, successfully attempted by Craig Venter and others at The Institute for Genomic Research, involves engineering existing prokaryotic cells with progressively fewer genes, attempting to discern at which point the most minimal requirements for life were reached. The biologist John Desmond Bernal coined the term Biopoesis for this process, and suggested that there were a number of clearly defined "stages" that could be recognised in explaining the origin of life.

- Stage 1: The origin of biological monomers
- Stage 2: The origin of biological polymers
- Stage 3: The evolution from molecules to cell

Bernal suggested that evolution may have commenced early, some time between Stage 1 and 2.

Origin of organic molecules

There are two possible sources of organic molecules on the early Earth:

1. Terrestrial origins–organic synthesis driven by impact shocks or by other energy sources (such as ultraviolet light or electrical discharges) (eg.Miller's experiments)
2. Extraterrestrial origins–delivery by objects (e.g. carbonaceous chondrites) or gravitational attraction of organic molecules or primitive life-forms from space

Recently, estimates of these sources suggest that the heavy bombardment before 3.5 Ga within the early atmosphere made available quantities of organics comparable to those produced by other energy sources.

"Soup" theory today: Miller's experiment and subsequent work

Biochemist Robert Shapiro has summarized the "Primordial Soup" theory of Oparin and Haldane in its "mature form" as follows:

1. The early Earth had a chemically reducing atmosphere.
2. This atmosphere, exposed to energy in various forms, produced simple organic compounds ("monomers").
3. These compounds accumulated in a "soup", which may have been concentrated at various locations (Shorelines, oceanic vents etc.).

4. By further transformation, more complex organic polymers— and ultimately life—developed in the soup.

Regarding the reducing atmosphere

Whether the mixture of gases used in the Miller–Urey experiment truly reflects the atmospheric content of early Earth is a controversial topic. Other less reducing gases produce a lower yield and variety. It was once thought that appreciable amounts of molecular oxygen were present in the prebiotic atmosphere, which would have essentially prevented the formation of organic molecules; however, the current scientific consensus is that such was not the case.

Regarding monomer formation

One of the most important pieces of experimental support for the "soup" theory came in 1953. A graduate student, Stanley Miller, and his professor, Harold Urey, performed an experiment that demonstrated how organic molecules could have spontaneously formed from inorganic precursors, under conditions like those posited by the Oparin-Haldane Hypothesis. The now-famous "Miller–Urey experiment" used a highly reduced mixture of gases–methane, ammonia and hydrogen–to form basic organic monomers, such as amino acids. This provided direct experimental support for the second point of the "soup" theory, and it is around the remaining two points of the theory that much of the debate now centers.

Apart from the Miller–Urey experiment, the next most important step in research on prebiotic organic synthesis was the demonstration by Joan Oró that the nucleic acid purine base, adenine, was formed by heating aqueous ammonium cyanide solutions. In support of abiogenesis in eutectic ice, more recent work demonstrated the formation of s-triazines (alternative nucleobases), pyrimidines (including cytosine and uracil), and adenine from urea solutions subjected to freeze-thaw cycles under a reductive atmosphere (with spark discharges as an energy source).

Regarding monomer accumulation

The "soup" theory relies on the assumption proposed by Darwin that in an environment with no pre-existing life, organic molecules may have accumulated and provided an environment for chemical evolution.

Regarding further transformation

The spontaneous formation of complex polymers from abiotically generated monomers under the conditions posited by the "soup" theory is not at all a straightforward process. Besides the necessary basic organic monomers, compounds that would have prohibited the formation of polymers were formed in high concentration during the Miller–Urey and Oró experiments. The Miller experiment, for example, produces many substances that would undergo cross-reactions with the amino acids or terminate the peptide chain.

More fundamentally, it can be argued that the most crucial challenge unanswered by this theory is how the relatively simple organic building blocks polymerise and form more complex structures, interacting in consistent ways to form a protocell. For example, in an aqueous environment hydrolysis of oligomers/polymers into their constituent monomers would be favored over the condensation of individual monomers into polymers.

The deep sea vent theory

The deep sea vent, or hydrothermal vent, theory for the origin of life on Earth posits that life may have begun at submarine hydrothermal vents, where hydrogen-rich fluids emerge from below the sea floor and interface with carbon dioxide-rich ocean water. Sustained chemical energy in such systems is derived from redox reactions, in which electron donors, such as molecular hydrogen, react with electron acceptors, such as carbon dioxide.

Fox's experiments

In the 1950s and 1960s, Sidney W. Fox studied the spontaneous formation of peptide structures under conditions that might plausibly have existed early in Earth's history. He demonstrated that amino acids could spontaneously form small peptides. These amino acids and small peptides could be encouraged to form closed spherical membranes, called protenoid microspheres, which show many of the basic characteristics of 'life'.

Eigen's hypothesis

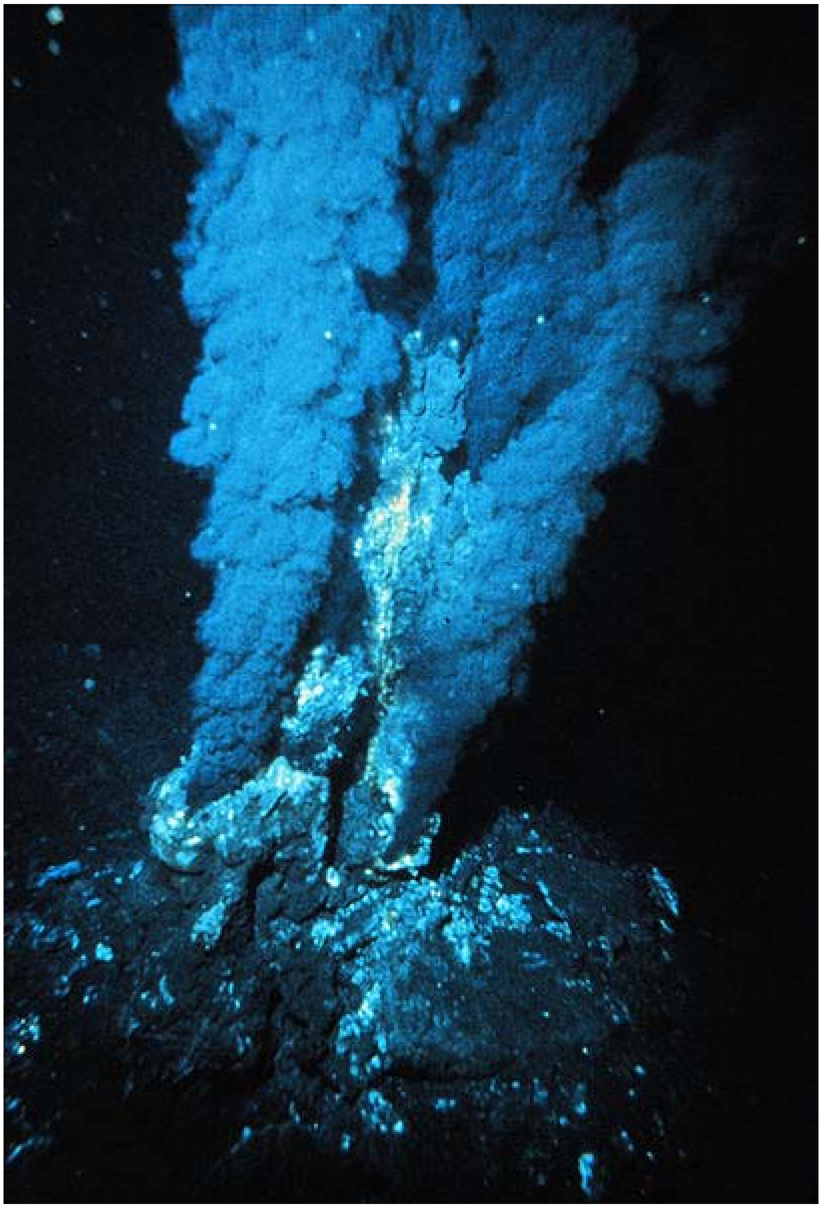
In the early 1970s the problem of the origin of life was approached by Manfred Eigen and Peter Schuster of the Max Planck Institute for Biophysical Chemistry. They examined the transient stages between the molecular chaos and a self-replicating hypercycle in a prebiotic soup.

In a hypercycle, the information storing system (possibly RNA) produces an enzyme, which catalyzes the formation of another information system, in sequence until the product of the last aids in the formation of the first information system. Mathematically treated, hypercycles could create quasispecies, which through natural selection entered into a form of Darwinian evolution. A boost to hypercycle theory was the discovery that RNA, in certain circumstances, forms itself into ribozymes, capable of catalyzing their own chemical reactions. However, these reactions are limited to self-excisions (in which a longer RNA molecule becomes shorter), and much rarer small additions that are incapable of coding for any useful protein. The hypercycle theory is further degraded since the hypothetical RNA would require the existence of complex biochemicals such as nucleotides which are not formed under the conditions proposed by the Miller–Urey experiment.

Hoffmann's contributions

Geoffrey W. Hoffmann, a student of Eigen, contributed to the concept of life involving both replication and metabolism emerging from catalytic noise. His contributions included showing that an early sloppy translation machinery can be stable against an error catastrophe of the type that had been envisaged as problematical by Leslie Orgel ("Orgel's paradox") and calculations regarding the occurrence of a set of required catalytic activities together with the exclusion of catalytic activities that would be disruptive. This is called the stochastic theory of the origin of life.

Wächtershäuser's hypothesis



Deep-sea black smoker

Another possible answer to this polymerization conundrum was provided in 1980s by the German chemist Günter Wächtershäuser, in his iron-sulfur world theory. In this theory, he postulated the evolution of (bio)chemical pathways as fundamentals of the evolution of life. Moreover, he presented a consistent system of tracing today's biochemistry back to ancestral reactions that provide alternative pathways to the synthesis of organic building blocks from simple gaseous compounds.

In contrast to the classical Miller experiments, which depend on external sources of energy (such as simulated lightning or UV irradiation), "Wächtershäuser systems" come with a built-in source of energy, sulfides of iron and other minerals (e.g. pyrite). The energy released from redox reactions of these metal sulfides is not only available for the synthesis of organic molecules, but also for the formation of oligomers and polymers. It is therefore hypothesized that such systems may be able to evolve into autocatalytic sets of self-replicating, metabolically active entities that would predate the life forms known today.

The experiment produced a relatively small yield of dipeptides (0.4% to 12.4%) and a smaller yield of tripeptides (0.10%) but the authors also noted that: "under these same conditions dipeptides hydrolysed rapidly."

Radioactive beach hypothesis

Zachary Adam at the University of Washington, Seattle, claims that stronger tidal processes from a much closer moon may have concentrated grains of uranium and other radioactive elements at the high water mark on primordial beaches where they may have been responsible for generating life's building blocks. According to computer models reported in *Astrobiology*, a deposit of such radioactive materials could show the same self-sustaining nuclear reaction as that found in the Oklo uranium ore seam in Gabon. Such radioactive beach sand provides sufficient energy to generate organic molecules, such as amino acids and sugars from acetonitrile in water. Radioactive monazite also releases soluble phosphate into regions between sand-grains, making it biologically "accessible". Thus amino acids, sugars and soluble phosphates can all be simultaneously produced, according to Adam. Radioactive actinides, then in greater concentrations, could have formed part of organo-metallic complexes. These complexes could have been important early catalysts to living processes.

John Parnell of the University of Aberdeen suggests that such a process could provide part of the "crucible of life" on any early wet rocky planet, so long as the planet is large enough to have generated a system of plate tectonics which brings radioactive minerals to the surface. As the early Earth is believed to have had many smaller "platelets" it would provide a suitable environment for such processes.

Thermodynamic Origin of Life: Ultraviolet and Temperature-Assisted Replication (UVTAR) Model

Karo Michaelian of the National Autonomous University of Mexico (UNAM) points out that any model for the origin of life must take into account the fact that life is an irreversible thermodynamic process which arises and persists to produce entropy. Entropy production is not incidental to the process of life, but rather the fundamental reason for its existence. Present day life augments the entropy production of Earth by catalysing the water cycle through evapotranspiration. Michaelian argues that if the thermodynamic function of life today is to produce entropy through coupling with the water cycle, then this probably was its function at its very beginnings. It turns out that both RNA and DNA when in water solution are very strong absorbers and extremely rapid dissipaters of ultraviolet light within the 200 nm - 300 nm wavelength range, just that high energy part of the sun's spectrum that could have penetrated the dense prebiotic atmosphere. Cnossen et al. have shown that the amount of UV light reaching the Earth's surface in the Archean could have been up to 31 orders of magnitude larger than it is today at 260 nm where RNA and DNA absorb most strongly. Absorption and dissipation of UV light by these organic molecules at the Archean ocean surface would have increased significantly the temperature of the surface skin layer leading to enhanced evaporation and thus augmenting the primitive water cycle. Since absorption and dissipation of high energy photons is an entropy producing process, Michaelian argues that non-equilbrium abiogenic synthesis of RNA and DNA utilizing UV light would have been thermodynamically favored.

A simple mechanism to explain the replication of RNA and DNA without the use of enzymes can also be given within the same thermodynamic framework by assuming that life arose when the temperature of the primitive seas had cooled to somewhat below the denaturing temperature of RNA or DNA (based on the ratio of $^{18}O/^{16}O$ found in cherts of the Barberton greenstone belt of South Africa of about 3.5 to 3.2 Ga., surface temperatures are predicted to have been around 70±15 °C, similar to RNA or DNA denaturing temperatures). During the night, the surface water temperature would be below the denaturing temperature and single strand RNA/DNA could act as a template for the formation of double strand RNA/DNA. During the daylight hours, RNA and DNA would absorb UV light and convert this directly to heating of the ocean surface, raising the local temperature enough to allow for denaturing of RNA and DNA. The copying process would be repeated during the cool period overnight. Such a temperature assisted mechanism of replication bears similarity to Polymerase Chain Reaction (PCR), a routine laboratory procedure to multiply DNA segments. Michaelian suggests that traditional origin of life research, expecting to describe the emergence of life from near-equilibrium conditions, is erroneous and that non-equilibrium conditions must be considered, in particular, the importance of entropy production to the emergence of life.

Since denaturation would be most probable in the late afternoon when the Archean sea surface temperature would be highest, and since late afternoon sunlight is somewhat circularly polarized, the homochirality of the organic molecules of life can also be explained within the proposed thermodynamic framework.

Models to explain homochirality

Some process in chemical evolution must account for the origin of homochirality, i.e. all building blocks in living organisms having the same "handedness" (amino acids being left-handed, nucleic acid sugars (ribose and deoxyribose) being right-handed, and chiral phosphoglycerides). Chiral molecules can be synthesized, but in the absence of a chiral source or a chiral catalyst, they are formed in a 50/50 mixture of both enantiomers. This is called a racemic mixture. Clark has suggested that homochirality may have started in space, as the studies of the amino acids on the Murchison meteorite showed L-alanine to be more than twice as frequent as its D form, and L-glutamic acid was more than 3 times prevalent than its D counterpart. It is suggested that polarised light has the power to destroy one enantiomer within the proto-planetary disk. Noyes showed that beta decay caused the breakdown of D-leucine, in a racemic mixture, and that the presence of ^{14}C, present in larger amounts in organic chemicals in the early Earth environment, could have been the cause. Robert M. Hazen reports upon experiments conducted in which various chiral crystal surfaces act as sites for possible concentration and assembly of chiral monomer units into macromolecules. Once established, chirality would be selected for. Work with organic compounds found on meteorites tends to suggest that chirality is a characteristic of abiogenic synthesis, as amino acids show a left-handed bias, whereas sugars show a predominantly right-handed bias.

Self-organization and replication

While features of self-organization and self-replication are often considered the hallmark of living systems, there are many instances of abiotic molecules exhibiting such characteristics under proper conditions. For example Martin and Russel show that physical compartmentation by cell membranes from the environment and self-organization of self-contained redox reactions are the most conserved attributes of living things, and they argue therefore that inorganic matter with such attributes would be life's most likely last common ancestor.

Virus self-assembly within host cells has implications for the study of the origin of life, as it lends further credence to the hypothesis that life could have started as self-assembling organic molecules.

From organic molecules to protocells

The question "How do simple organic molecules form a protocell?" is largely unanswered but there are many hypotheses. Some of these postulate the early appearance of nucleic acids ("genes-first") whereas others postulate the evolution of biochemical reactions and pathways first ("metabolism-first"). Recently, trends are emerging to create hybrid models that combine aspects of both.

"Genes first" models: the RNA world

The RNA world hypothesis describes an early Earth with self-replicating and catalytic RNA but no DNA or proteins. This has spurred scientists to try to determine if relatively short RNA molecules could have spontaneously formed that were capable of catalyzing their own continuing replication. A number of hypotheses of modes of formation have been put forward. Early cell membranes could have formed spontaneously from proteinoids, protein-like molecules that are produced when amino acid solutions are heated–when present at the correct concentration in aqueous solution, these form microspheres which are observed to behave similarly to membrane-enclosed compartments. Other possibilities include systems of chemical reactions taking place within clay substrates or on the surface of pyrite rocks. Factors supportive of an important role for RNA in early life include its ability to act both to store information and catalyse chemical reactions (as a ribozyme); its many important roles as an intermediate in the expression and maintenance of the genetic information (in the form of DNA) in modern organisms; and the ease of chemical synthesis of at least the components of the molecule under conditions approximating the early Earth. Relatively short RNA molecules which can duplicate others have been artificially produced in the lab. Such replicase RNA, which functions as both code and catalyst provides a template upon which copying can occur. Jack Szostak has shown that certain catalytic RNAs can, indeed, join smaller RNA sequences together, creating the potential, in the right conditions for self-replication. If these were present, Darwinian selection would favour the proliferation of such self-catalysing structures, to which further functionalities could be added. Lincoln and Joyce identified an RNA enzyme capable of self sustained replication.

Researchers have pointed out difficulties for the abiotic synthesis of nucleotides from cytosine and uracil. Cytosine has a half-life of 19 days at 100 °C (212 °F) and 17,000 years in freezing water. Larralde et al., say that "the generally accepted prebiotic synthesis of ribose, the formose reaction, yields numerous sugars without any selectivity." and they conclude that their "results suggest that the backbone of the first genetic material could not have contained ribose or other sugars because of their instability." The ester linkage of ribose and phosphoric acid in RNA is known to be prone to hydrolysis.

A slightly different version of the RNA-world hypothesis is that a different type of nucleic acid, such as PNA, TNA or GNA, was the first one to emerge as a self-reproducing molecule, to be replaced by RNA only later. Pyrimidine ribonucleosides and their respective nucleotides have been prebiotically synthesised by a sequence of reactions which by-pass the free sugars, and are assembled in a stepwise fashion by going against the dogma that nitrogenous and oxygenous chemistries should be avoided. In a series of publications, The Sutherland Group at the School of Chemistry, University of Manchester have demonstrated high yielding routes to cytidine and uridine ribonucleotides built from small 2 and 3 carbon fragments such as glycolaldehyde, glyceraldehyde or glyceraldehyde-3-phosphate, cyanamide and cyanoacetylene. One of the steps in this sequence allows the isolation of enantiopure ribose aminooxazoline if the enantiomeric excess of glyceraldehyde is 60 % or greater. This can be viewed as a prebiotic purification step, where the said compound spontaneously crystallised out from a mixture of the other pentose aminooxazolines. Ribose aminooxazoline can then react with cyanoacetylene in a mild and highly efficient manner to give the alpha cytidine ribonucleotide. Photoanomerization with UV light allows for inversion about the 1' anomeric centre to give the correct beta stereochemistry. In 2009 they showed that the same simple building blocks allow access, via phosphate controlled nucleobase elaboration, to 2',3'-cyclic pyrimidine nucleotides directly, which are known to be able to polymerise into RNA. This paper also highlights the possibility for the photo-sanitization of the pyrimidine-2',3'-cyclic phosphates. James Ferris's studies have shown that clay minerals of montmorillonite will catalyze the formation of RNA in aqueous solution, by joining activated mono RNA nucleotides to join together to form longer chains. Although these chains have random sequences, the possibility that one sequence began to non-randomly increase its frequency by increasing the speed of its catalysis is possible to "kick start" biochemical evolution.

"Metabolism first" models

Several models reject the idea of the self-replication of a "naked-gene" and postulate the emergence of a primitive metabolism which could provide an environment for the later emergence of RNA replication.

Iron-sulfur world

One of the earliest incarnations of this idea was put forward in 1924 with Alexander Oparin's notion of primitive self-replicating vesicles which predated the discovery of the structure of DNA. More recent variants in the 1980s and 1990s include Günter Wächtershäuser's iron-sulfur world theory and models introduced by Christian de Duve based on the chemistry of thioesters. More abstract and theoretical arguments for the plausibility of the emergence of metabolism without the presence of genes include a mathematical model introduced by Freeman Dyson in the early 1980s and Stuart Kauffman's notion of collectively autocatalytic sets, discussed later in that decade.

However, the idea that a closed metabolic cycle, such as the reductive citric acid cycle, could form spontaneously (proposed by Günter Wächtershäuser) remains debated. In an article entitled "Self-Organizing Biochemical Cycles", the late Leslie Orgel summarized his analysis of the proposal by stating, "There is at present no reason to expect that multistep cycles such as the reductive citric acid cycle will self-organize on the surface of FeS/FeS2 or some other mineral." It is possible that another type of metabolic pathway was used at the beginning of life. For example, instead of the reductive citric acid cycle, the "open" acetyl-CoA pathway (another one of the five recognised ways of carbon dioxide fixation in nature today) would be compatible with the idea of self-organisation on a metal sulfide surface. The key enzyme of this pathway, carbon monoxide dehydrogenase/acetyl-CoA synthase harbours mixed nickel-iron-sulfur clusters in its reaction centers and catalyses the formation of acetyl-CoA (which may be regarded as a modern form of acetyl-thiol) in a single step.

Thermosynthesis world

Today's bioenergetic process of fermentation is related to the just mentioned citric acid cycle or the Acetyl-CoA pathway that have been connected to the primordial iron-sulfur world. In a different approach, today's bioenergetic process of chemiosmosis, which plays an essential role in cellular respiration and photosynthesis, is considered as more fundamental than fermentation: in Anthonie Muller's "thermosynthesis world" the ATP Synthase enzyme that sustains chemiosmosis is proposed as today's enzyme that is the closest connected to the first metabolic process.

First life needed an energy source to bring about the condensation reaction that yielded the peptide bonds of proteins and the phosphodiester bonds of RNA. In a generalization and thermal variation of the binding change mechanism of today's ATP Synthase, the "First Protein" would have bound substrates (peptides, phosphate, nucleosides, RNA 'monomers') and condensed them to a reaction product that remained bound until it after a temperature change was released upon a thermal unfolding.

The energy source of the thermosynthesis world was thermal cycling, the result of suspension of the protocell in a convection current, as is plausible in a volcanic hot spring; the convection accounts for the self-organization and dissipative structure required in any origin of life model. The still ubiquitous role of thermal cycling in germination and cell division is considered a relic of primordial thermosynthesis.

By phosphorylating cell membrane lipids, this 'First Protein' gave a selective advantage to the lipid protocell that contained the protein. In the beginning this First Protein also synthesized a library with many proteins, of which only a minute fraction had thermosynthesis capabilities. Just as proposed by Dyson for the first proteins, the First Protein propagated functionally: it made daughters with similar capabilities, but it did not copy itself. Functioning daughters consisted of different amino acid sequences.

Over a long time, RNA sequences were selected among the at first randomly synthesized RNAs by the criterion of speed and efficiency increase of First Protein synthesis, for instance by the creation of RNA that functioned as messenger RNA, Transfer RNA and ribosomal RNA, or, even more generally, all the components of the RNA World were also generated and selected. The thermosynthesis world therefore in theory accounts for the origin of the genetic machinery.

Whereas the iron-sulfur world identifies a circular pathway as the most simple—and therefore assumes the existence of enzymes—the thermosynthesis world does not even invoke a pathway, and does not assume the existence of regular enzymes: ATP Synthase's binding change mechanism resembles a physical adsorption process that yields free energy, rather than a regular enzyme's mechanism, which decreases the free energy. The RNA World also implies the existence of several enzymes. But even the emergence of a single enzyme by chance is implausible. The thermosynthesis world is therefore more simple, and thus more plausible, than the iron-sulfur and RNA worlds.

Possible role of bubbles

Waves breaking on the shore create a delicate foam composed of bubbles. Winds sweeping across the ocean have a tendency to drive things to shore, much like driftwood collecting on the beach. It is possible that organic molecules were concentrated on the shorelines in much the same way.

Shallow coastal waters also tend to be warmer, further concentrating the molecules through evaporation. While bubbles composed mostly of water burst quickly, water containing amphiphiles forms much more stable bubbles, lending more time to the particular bubble to perform these crucial reactions.

Amphiphiles are oily compounds containing a hydrophilic head on one or both ends of a hydrophobic molecule. Some amphiphiles have the tendency to spontaneously form membranes in water. A spherically closed membrane contains water and is a hypothetical precursor to the modern cell membrane. If a protein would increase the integrity of its parent bubble, that bubble had an advantage, and was placed at the top of the natural selection waiting list. Primitive reproduction can be envisioned when the bubbles burst, releasing the results of the 'experiment' into the surrounding medium. Once enough of the 'right stuff' was released into the medium, the development of the first prokaryotes, eukaryotes, and multicellular organisms could be achieved.

Similarly, bubbles formed entirely out of protein-like molecules, called microspheres, will form spontaneously under the right conditions. But they are not a likely precursor to the modern cell membrane, as cell membranes are composed primarily of lipid compounds rather than amino-acid compounds.

A recent model by Fernando and Rowe suggests that the enclosure of an autocatalytic non-enzymatic metabolism within protocells may have been one way of avoiding the side-reaction problem that is typical of metabolism first models.

Other models

Autocatalysis

In 1993 Stuart Kauffman proposed that life initially arose as autocatalytic chemical networks.

British ethologist Richard Dawkins wrote about autocatalysis as a potential explanation for the origin of life in his 2004 book *The Ancestor's Tale*. Autocatalysts are substances which catalyze the production of themselves, and therefore have the property of being a simple molecular replicator. In his book, Dawkins cites experiments performed by Julius Rebek and his colleagues at the Scripps Research Institute in California in which they combined amino adenosine and pentafluorophenyl ester with the autocatalyst amino adenosine triacid ester (AATE). One system from the experiment contained variants of AATE which catalysed the synthesis of themselves. This experiment demonstrated the possibility that autocatalysts could exhibit competition within a population of entities with heredity, which could be interpreted as a rudimentary form of natural selection.

Clay theory

A model for the origin of life based on clay was forwarded by A. Graham Cairns-Smith of the University of Glasgow in 1985 and explored as a plausible illustration by several other scientists, including Richard Dawkins. Clay theory postulates that complex organic molecules arose gradually on a pre-existing, non-organic replication platform—silicate crystals in solution. Complexity in companion molecules developed as a function of selection pressures on types of clay crystal is then exapted to serve the replication of organic molecules independently of their silicate "launch stage".

Cairns-Smith is a staunch critic of other models of chemical evolution. However, he admits that like many models of the origin of life, his own also has its shortcomings (Horgan 1991).

In 2007, Kahr and colleagues reported their experiments to examine the idea that crystals can act as a source of transferable information, using crystals of potassium hydrogen phthalate. "Mother" crystals with imperfections were cleaved and used as seeds to grow "daughter" crystals from solution. They then examined the distribution of imperfections in the crystal system and found that the imperfections in the mother crystals were indeed reproduced in the daughters. The daughter crystals had many additional imperfections. For a gene-like behavior the additional imperfections should be much less than the parent ones, thus Kahr concludes that the crystals "were not faithful enough to store and transfer information from one generation to the next".

Gold's "Deep-hot biosphere" model

In the 1970s, Thomas Gold proposed the theory that life first developed not on the surface of the Earth, but several kilometers below the surface. The discovery in the late 1990s of nanobes (filamental structures that are smaller than bacteria, but that may contain DNA) in deep rocks might be seen as lending support to Gold's theory.

It is now reasonably well established that microbial life is plentiful at shallow depths in the Earth, up to 5 kilometres (3.1 mi) below the surface, in the form of extremophile archaea, rather than the better-known eubacteria (which live in more accessible conditions). It is claimed that discovery of microbial life below the surface of another body in our solar system would lend significant credence to this theory. Thomas Gold also asserted that a trickle of food from a deep, unreachable, source is needed for survival because life arising in a puddle of organic material is likely to consume all of its food and become extinct. Gold's theory is that flow of food is due to out-gassing of primordial methane from the Earth's mantle; more conventional explanations of the food supply of deep microbes (away from sedimentary carbon compounds) is that the organisms subsist on hydrogen released by an interaction between water and (reduced) iron compounds in rocks.

"Primitive" extraterrestrial life

An alternative to Earthly abiogenesis is the hypothesis that primitive life may have originally formed extraterrestrially, either in space or on a nearby planet (Mars). (Note that exogenesis is related to, but not the same as, the notion of panspermia). A supporter of this theory was Francis Crick.

Organic compounds are relatively common in space, especially in the outer solar system where volatiles are not evaporated by solar heating. Comets are encrusted by outer layers of dark material, thought to be a tar-like substance composed of complex organic material formed from simple carbon compounds after reactions initiated mostly by irradiation by ultraviolet light. It is supposed that a rain of material from comets could have brought significant quantities of such complex organic molecules to Earth.

An alternative but related hypothesis, proposed to explain the presence of life on Earth so soon after the planet had cooled down, with apparently very little time for prebiotic evolution, is that life formed first on early Mars. Due to its smaller size Mars cooled before Earth (a difference of hundreds of millions of years), allowing prebiotic processes there while Earth was still too hot. Life was then transported to the cooled Earth when crustal material was blasted off Mars by asteroid and comet impacts. Mars continued to cool faster and eventually became hostile to the continued evolution or even existence of life (it lost its atmosphere due to low volcanism); Earth is following the same fate as Mars, but at a slower rate.

Neither hypothesis actually answers the question of how life first originated, but merely shifts it to another planet or a comet. However, the advantage of an extraterrestrial origin of primitive life is that life is not required to have evolved on each planet it occurs on, but rather in a single location, and then spread about the galaxy to other star systems via cometary and/or meteorite impact. Evidence to support the plausibility of the concept is scant, but it finds support in recent study of Martian meteorites found in Antarctica and in studies of extremophile microbes. Additional support comes from a recent discovery of a bacterial ecosytem whose energy source is radioactivity.

A 2001 experiment led by Jason Dworkin subjected a frozen mixture of water, methanol, ammonia and carbon monoxide to UV radiation, mimicking conditions found in an extraterrestrial environment. This combination yielded large amounts of organic material that self-organised to form bubbles or micelles when immersed in water. Dworkin considered these bubbles to resemble cell membranes that enclose and concentrate the chemistry of life, separating their interior from the outside world.

The bubbles produced in these experiments were between 10 to 40 micrometres (0.00039 to 0.0016 in), or about the size of red blood cells. Remarkably, the bubbles fluoresced, or glowed, when exposed to UV light. Absorbing UV and converting it into visible light in this way was considered one possible way of providing energy to a primitive cell. If such bubbles played a role in the origin of life, the fluorescence could have been a precursor to primitive photosynthesis. Such fluorescence also provides the benefit of acting as a sunscreen, diffusing any damage that otherwise would be inflicted by UV radiation. Such a protective function would have been vital for life on the early Earth, since the ozone layer, which blocks out the sun's most destructive UV rays, did not form until after photosynthetic life began to produce oxygen.

Extraterrestrial amino acids

Another idea is that amino acids which were formed extra-terrestrially arrived on Earth via comets. In 2009 it was announced by NASA that scientists have identified one of the fundamental chemical buildings blocks of life in a comet for the first time: glycine, an amino acid, was detected in the material ejected from Comet Wild-2 in 2004 and grabbed by NASA's Stardust probe. Tiny grains, just a few thousandths of a millimetre in size, were collected from the comet and returned to Earth in 2006 in a sealed capsule, and distributed among the world's leading astro-biology labs. NASA said in a statement that it took some time for the investigating team, led by Dr Jamie Elsila, to convince itself that the glycine signature found in Stardust's sample bay was genuine and not just Earthly contamination. Glycine has been detected in meteorites before and there are also observations in interstellar gas clouds claimed for telescopes, but the Stardust find is described as a first in cometary material. It is known that prior to the emergence of life on Earth, the early solar system's planets were regularly bombarded by comets. Dr. Carl Pilcher, who leads NASA's Astrobiology Institute commented that "The discovery of glycine in a comet supports the idea that the fundamental building blocks of life are prevalent in space, and strengthens the argument that life in the Universe may be common rather than rare."

Lipid World

This theory postulates that the first self-replicating object was lipid-like. It is known that phospholipids form bilayers in water while under agitation– the same structure as in cell membranes. These molecules were not present on early Earth, however other amphiphilic long chain molecules also form membranes. Furthermore, these bodies may expand (by insertion of additional lipids), and under excessive expansion may undergo spontaneous splitting which preserves the same size and composition of lipids in the two progenies. The main idea in this theory is that the molecular composition of the lipid bodies is the preliminary way for information storage, and evolution led to the appearance of polymer entities such as RNA or DNA that may store information favorably. Still, no biochemical mechanism has been offered to support the Lipid World theory.

Polyphosphates

The problem with most scenarios of abiogenesis is that the thermodynamic equilibrium of amino acid versus peptides is in the direction of separate amino acids. What has been missing is some force that drives polymerization. The resolution of this problem may well be in the properties of polyphosphates. Polyphosphates are formed by polymerization of ordinary monophosphate ions PO_4^{-3}. Several mechanisms for such polymerization have been suggested. Polyphosphates cause polymerization of amino acids into peptides. They are also logical precursors in the synthesis of such key biochemical compounds as ATP. A key issue seems to be that calcium reacts with soluble phosphate to form insoluble calcium phosphate (apatite), so some plausible mechanism must be found to keep calcium ions from causing precipitation of phosphate. There has been much work on this topic over the years, but an interesting new idea is that meteorites may have introduced reactive phosphorus species on the early Earth.

PAH world hypothesis

Other sources of complex molecules have been postulated, including extraterrestrial stellar or interstellar origin. For example, from spectral analyses, organic molecules are known to be present in comets and meteorites. In 2004, a team detected traces of polycyclic aromatic hydrocarbons (PAH's) in a nebula. Those are the most complex molecules so far found in space. The use of PAH's has also been proposed as a precursor to the RNA world in the PAH world hypothesis. The Spitzer Space Telescope has recently detected a star, HH 46-IR, which is forming by a process similar to that by which the sun formed. In the disk of material surrounding the star, there is a very large range of molecules, including cyanide compounds, hydrocarbons, and carbon monoxide. PAHs have also been found all over the surface of galaxy M81, which is 12 million light years away from the Earth, confirming their widespread distribution in space.

Multiple genesis

Different forms of life may have appeared quasi-simultaneously in the early history of Earth. The other forms may be extinct, leaving distinctive fossils through their different biochemistry (e.g., using arsenic instead of phosphorus), survive as extremophiles, or simply be unnoticed through their being analogous to organisms of the current life tree. Hartman for example combines a number of theories together, by proposing that:

The first organisms were self-replicating iron-rich clays which fixed carbon dioxide into oxalic and other dicarboxylic acids. This system of replicating clays and their metabolic phenotype then evolved into the sulfide rich region of the hotspring acquiring the ability to fix nitrogen. Finally phosphate was incorporated into the evolving system which allowed the synthesis of nucleotides and phospholipids. If biosynthesis recapitulates biopoesis, then the synthesis of amino acids preceded the synthesis of the purine and pyrimidine bases. Furthermore the polymerization of the amino acid thioesters into polypeptides preceded the directed polymerization of amino acid esters by polynucleotides.

Lynn Margulis's endosymbiotic theory suggests that multiple forms of bacteria entered into symbiotic relationship to form the eukaryotic cell. The horizontal transfer of genetic material between bacteria promotes such symbiotic relationships, and thus many separate organisms may have contributed to building what has been recognised as the Last Universal Common Ancestor (LUCA) of modern organisms. James Lovelock's Gaia theory, proposes that such bacterial symbiosis establishes the environment as a system produced by and supportive of life. His arguments strongly weaken the case for life having evolved elsewhere in the solar system.

Chapter- 6

Human Evolution

Human evolution, or ***anthropogenesis***, is the origin and evolution of *Homo sapiens* as a distinct species from other hominids, great apes and placental mammals. The study of human evolution encompasses many scientific disciplines, including physical anthropology, primatology, archaeology, linguistics and genetics.

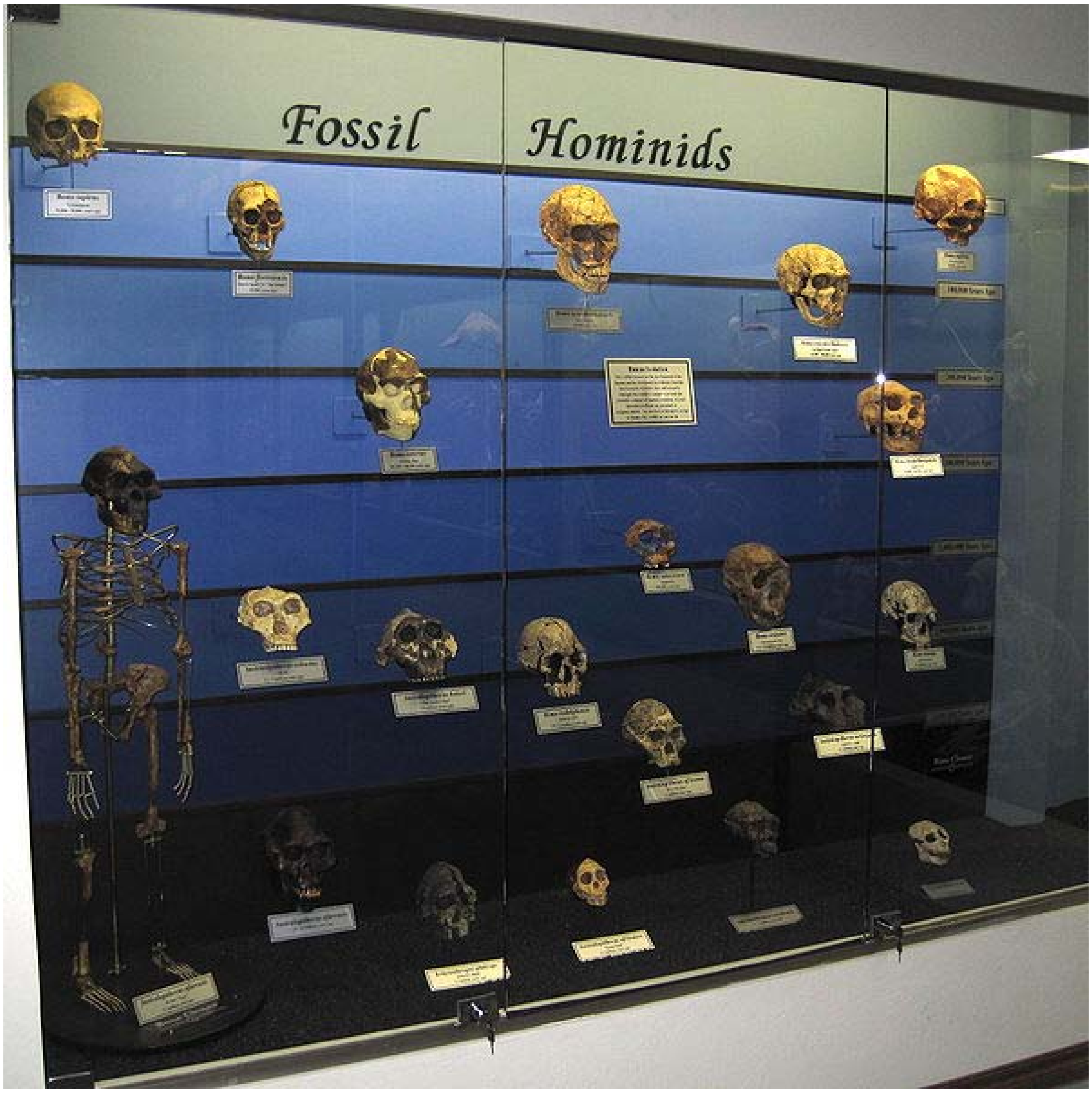
The term "human" in the context of human evolution refers to the genus *Homo*, but studies of human evolution usually include other hominids, such as the Australopithecines, from which the genus *Homo* had diverged by about 2.3 to 2.4 million years ago in Africa. Scientists have estimated that humans branched off from their common ancestor with chimpanzees about 5–7 million years ago. Several species and subspecies of *Homo* evolved and are now extinct. These include *Homo erectus*, which inhabited Asia, and *Homo sapiens neanderthalensis*, which inhabited Europe. Archaic *Homo sapiens* evolved between 400,000 and 250,000 years ago.

The dominant view among scientists concerning the origin of anatomically modern humans is the "Out of Africa" or recent African origin hypothesis, which argues that *Homo sapiens* arose in Africa and migrated out of the continent around 50,000 to 100,000 years ago, replacing populations of *Homo erectus* in Asia and *Homo neanderthalensis* in Europe. Scientists supporting the alternative multiregional hypothesis argue that *Homo sapiens* evolved as geographically separate but interbreeding populations stemming from a worldwide migration of *Homo erectus* out of Africa nearly 2.5 million years ago.

History of ideas

The word *homo*, the name of the biological genus to which humans belong, is Latin for "human". It was chosen originally by Carolus Linnaeus in his classification system. The word "human" is from the Latin *humanus*, the adjectival form of *homo*. The Latin "homo" derives from the Indo-European root, *dhghem*, or "earth".

Carolus Linnaeus and other scientists of his time also considered the great apes to be the closest relatives of humans due to morphological and anatomical similarities. The possibility of linking humans with earlier apes by descent only became clear after 1859 with the publication of Charles Darwin's *On the Origin of Species*. This argued for the idea of the evolution of new species from earlier ones. Darwin's book did not address the question of human evolution, saying only that "Light will be thrown on the origin of man and his history".



Fossil Hominid Evolution Display at The Museum of Osteology, Oklahoma City, USA

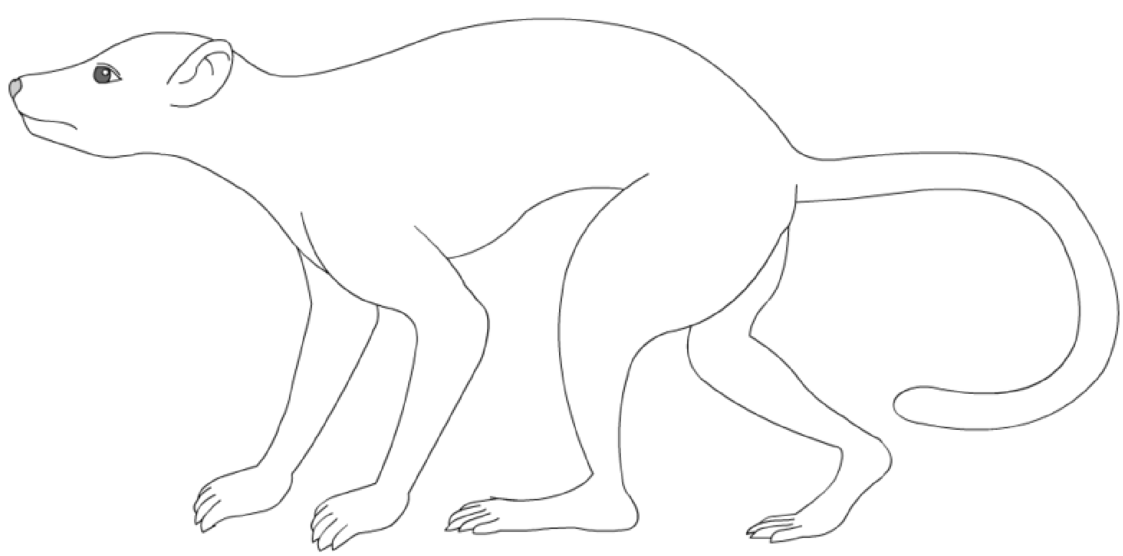
The first debates about the nature of human evolution arose between Thomas Huxley and Richard Owen. Huxley argued for human evolution from apes by illustrating many of the similarities and differences between humans and apes and did so particularly in his 1863 book *Evidence as to Man's Place in Nature*. However, many of Darwin's early supporters (such as Alfred Russel Wallace and Charles Lyell) did not agree that the origin of the mental capacities and the moral sensibilities of humans could be explained by natural selection. Darwin applied the theory of evolution and sexual selection to humans when he published *The Descent of Man* in 1871.

A major problem was the lack of fossil intermediaries. It was only in the 1920s that such fossils were discovered in Africa. In 1925, Raymond Dart described *Australopithecus africanus*. The type specimen was the Taung Child, an Australopithecine infant discovered in a cave. The child's remains were a remarkably well-preserved tiny skull and an endocranial cast of the individual's brain. Although the brain was small (410 cm^3), its shape was rounded, unlike that of chimpanzees and gorillas, and more like a modern human brain. Also, the specimen showed short canine teeth, and the position of the foramen magnum was evidence of bipedal locomotion. All of these traits convinced Dart that the Taung baby was a bipedal human ancestor, a transitional form between apes and humans.

The classification of humans and their relatives has changed considerably over time. The gracile Australopithecines are now thought to be ancestors of the genus *Homo*, the group to which modern humans belong. Both Australopithecines and *Homo sapiens* are part of the tribe Hominini. Recent data suggests Australopithecines were a diverse group and that *A. africanus* may not be a direct ancestor of modern humans. Reclassification of Australopithecines that originally were split into either gracile or robust varieties has put the latter into a family of its own, *Paranthropus*. Taxonomists place humans, Australopithecines and related species in the same family as other great apes, in the Hominidae.

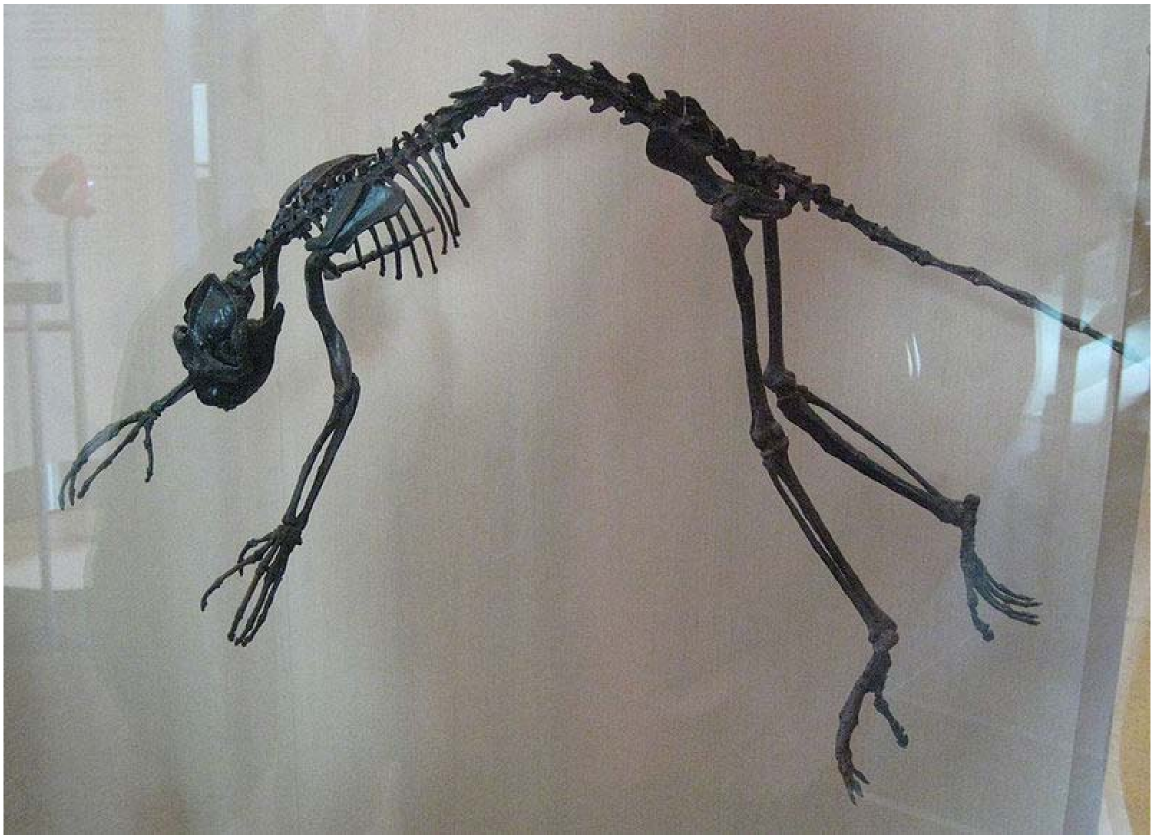
Before *Homo*

Evolution of the great apes



Plesiadapis

The evolutionary history of the primates can be traced back 65 million years, as one of the oldest of all surviving placental mammal groups. The oldest known primate-like mammal species, the Plesiadapis, come from North America, but they were widespread in Eurasia and Africa during the tropical conditions of the Paleocene and Eocene.



Notharctus

With the beginning of modern climates, marked by the formation of the first Antarctic ice in the early Oligocene around 30 million years ago. A primate from this time was *Notharctus*. Fossil evidence found in Germany in the 1980s was determined to be about 16.5 million years old, some 1.5 million years older than similar species from East Africa and challenging the original theory regarding human ancestry originating on the African continent.

David Begun says that these primates flourished in Eurasia and that the lineage leading to the African apes and humans— including Dryopithecus—migrated south from Europe or Western Asia into Africa. The surviving tropical population, which is seen most completely in the upper Eocene and lowermost Oligocene fossil beds of the Fayum depression southwest of Cairo, gave rise to all living primates—lemurs of Madagascar, lorises of Southeast Asia, galagos or "bush babies" of Africa, and the anthropoids; platyrrhines or New World monkeys, and catarrhines or Old World monkeys and the great apes and humans.

The earliest known catarrhine is *Kamoyapithecus* from uppermost Oligocene at Eragaleit in the northern Kenya Rift Valley, dated to 24 million years ago. Its ancestry is generally thought to be species related to *Aegyptopithecus*, *Propliopithecus*, and *Parapithecus* from the Fayum, at around 35 million years ago. In 2010, *Saadanius* was described as a close relative of the last common ancestor of the crown catarrhines, and tentatively dated to 29–28 million years ago, helping to fill an 11-million-year gap in the fossil record.

Reconstructed tailless *Proconsul* skeleton

In the early Miocene, about 22 million years ago, the many kinds of arboreally adapted primitive catarrhines from East Africa suggest a long history of prior diversification. Fossils at 20 million years ago include fragments attributed to *Victoriapithecus*, the earliest Old World Monkey. Among the genera thought to be in the ape lineage leading up to 13 million years ago are *Proconsul*, *Rangwapithecus*, *Dendropithecus*, *Limnopithecus*, *Nacholapithecus*, *Equatorius*, *Nyanzapithecus*, *Afropithecus*, *Heliopithecus*, and *Kenyapithecus*, all from East Africa. The presence of other generalized non-cercopithecids of middle Miocene age from sites far distant—*Otavipithecus* from cave deposits in Namibia, and *Pierolapithecus* and *Dryopithecus* from France, Spain and Austria—is evidence of a wide diversity of forms across Africa and the Mediterranean basin during the relatively warm and equable climatic regimes of the early and middle Miocene. The youngest of the Miocene hominoids, *Oreopithecus*, is from 9 million year old coal beds in Italy.

Molecular evidence indicates that the lineage of gibbons (family Hylobatidae) became distinct from Great Apes between 18 and 12 million years ago, and that of orangutans (subfamily Ponginae) became distinct from the other Great Apes at about 12 million years; there are no fossils that clearly document the ancestry of gibbons, which may have originated in a so-far-unknown South East Asian hominoid population, but fossil proto-orangutans may be represented by *Ramapithecus* from India and *Griphopithecus* from Turkey, dated to around 10 million years ago.

Divergence of the human lineage from other Great Apes

Species close to the last common ancestor of gorillas, chimpanzees and humans may be represented by *Nakalipithecus* fossils found in Kenya and *Ouranopithecus* found in Greece. Molecular evidence suggests that between 8 and 4 million years ago, first the gorillas, and then the chimpanzees (genus *Pan*) split off from the line leading to the humans; human DNA is approximately 98.4% identical to that of chimpanzees when comparing single nucleotide polymorphisms. The fossil record of gorillas and chimpanzees is quite limited. Both poor preservation (rain forest soils tend to be acidic and dissolve bone) and sampling bias probably contribute to this problem.

Other hominines likely adapted to the drier environments outside the equatorial belt, along with antelopes, hyenas, dogs, pigs, elephants, and horses. The equatorial belt contracted after about 8 million years ago. Fossils of these hominans - the species in the human lineage following divergence from the chimpanzees - are relatively well known. The earliest are *Sahelanthropus tchadensis* (7 Ma) and *Orrorin tugenensis* (6 Ma), followed by:

- *Ardipithecus* (5.5–4.4 Ma), with species *Ar. kadabba* and *Ar. ramidus*;
- *Australopithecus* (4–1.8 Ma), with species *Au. anamensis*, *Au. afarensis*, *Au. africanus*, *Au. bahrelghazali*, *Au. garhi*, and *Au. sediba*;
- *Kenyanthropus* (3–2.7 Ma), with species *Kenyanthropus platyops*;
- *Paranthropus* (3–1.2 Ma), with species *P. aethiopicus*, *P. boisei*, and *P. robustus*;
- *Homo* (2 Ma–present), with species *Homo habilis*, *Homo rudolfensis*, *Homo ergaster*, *Homo georgicus*, *Homo antecessor*, *Homo cepranensis*, *Homo erectus*, *Homo heidelbergensis*, *Homo rhodesiensis*, *Homo neanderthalensis*, *Homo sapiens idaltu*, *Archaic Homo sapiens*, *Homo floresiensis*.

Genus *Homo*

Homo sapiens is the only extant species of its genus, *Homo*. While some other, extinct, *Homo* species might have been ancestors of *Homo sapiens*, many were likely our "cousins", having speciated away from our ancestral line. There is not yet a consensus as to which of these groups should count as separate species and which as subspecies. In some cases this is due to the dearth of fossils, in other cases it is due to the slight differences used to classify species in the *Homo* genus. The Sahara pump theory (describing an occasionally passable "wet" Sahara Desert) provides an explanation of the early variation in the genus *Homo*.

Based on archaeological and paleontological evidence, it has been possible to infer, to some extent, the ancient dietary practices of various *Homo* species and to study the role of diet in physical and behavioral evolution within *Homo*.

H. habilis

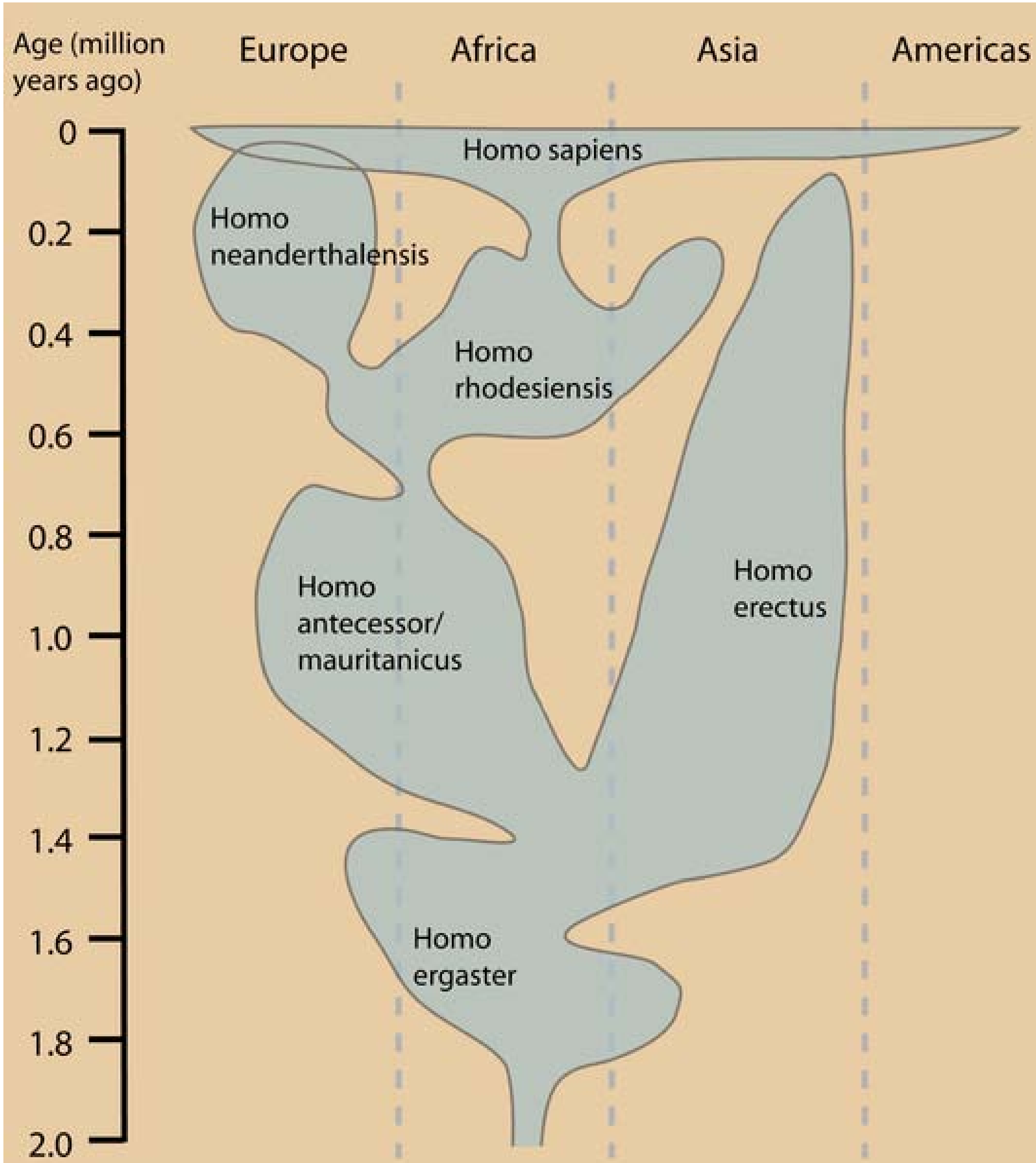
Homo habilis lived from about 2.4 to 1.4 Ma. *Homo habilis*, the first species of the genus *Homo*, evolved in South and East Africa in the late Pliocene or early Pleistocene, 2.5–2 Ma, when it diverged from the Australopithecines. *Homo habilis* had smaller molars and larger brains than the Australopithecines, and made tools from stone and perhaps animal bones. One of the first known hominids, it was nicknamed 'handy man' by its discoverer, Louis Leakey due to its association with stone tools. Some scientists have proposed moving this species out of *Homo* and into *Australopithecus* due to the morphology of its skeleton being more adapted to living on trees rather than to moving on two legs like *Homo sapiens*.

H. rudolfensis and *H. georgicus*

These are proposed species names for fossils from about 1.9–1.6 Ma, the relation of which with *Homo habilis* is not yet clear.

- *Homo rudolfensis* refers to a single, incomplete skull from Kenya. Scientists have suggested that this was another *Homo habilis*, but this has not been confirmed.
- *Homo georgicus*, from Georgia, may be an intermediate form between *Homo habilis* and *Homo erectus*, or a sub-species of *Homo erectus*.

H. ergaster and *H. erectus*



One current view of the temporal and geographical distribution of hominid populations. Other interpretations differ mainly in the taxonomy and geographical distribution of hominid species.

The first fossils of *Homo erectus* were discovered by Dutch physician Eugene Dubois in 1891 on the Indonesian island of Java. He originally gave the material the name *Pithecanthropus erectus* based on its morphology that he considered to be intermediate between that of humans and apes. *Homo erectus* (H erectus) lived from about 1.8 Ma to about 70,000 years ago (which would indicate that they were probably wiped out by the Toba catastrophe; however, *Homo erectus soloensis* and *Homo floresiensis* survived it). Often the early phase, from 1.8 to 1.25 Ma, is considered to be a separate species, *Homo ergaster*, or it is seen as a subspecies of *Homo erectus*, *Homo erectus ergaster*. In the early Pleistocene, 1.5–1 Ma, in Africa, Asia, and Europe, some populations of *Homo habilis* are thought to have evolved larger brains and made more elaborate stone tools; these differences and others are sufficient for anthropologists to classify them as a new species, *Homo erectus*. In addition *Homo erectus* was the first human ancestor to walk truly upright. This was made possible by the evolution of locking knees and a different location of the foramen magnum (the hole in the skull where the spine enters). They may have used fire to cook their meat.

A famous example of *Homo erectus* is Peking Man; others were found in Asia (notably in Indonesia), Africa, and Europe. Many paleoanthropologists now use the term *Homo ergaster* for the non-Asian forms of this group, and reserve *Homo erectus* only for those fossils that are found in Asia and meet certain skeletal and dental requirements which differ slightly from *H. ergaster*.

H. cepranensis and *H. antecessor*

These are proposed as species that may be intermediate between *H. erectus* and *H. heidelbergensis*.

- *H. antecessor* is known from fossils from Spain and England that are dated 1.2 Ma–500 ka.
- *H. cepranensis* refers to a single skull cap from Italy, estimated to be about 800,000 years old.

H. heidelbergensis

H. heidelbergensis (Heidelberg Man) lived from about 800,000 to about 300,000 years ago. Also proposed as *Homo sapiens heidelbergensis* or *Homo sapiens paleohungaricus*.

H. rhodesiensis, and the Gawis cranium

- *H. rhodesiensis*, estimated to be 300,000–125,000 years old. Most current experts believe Rhodesian Man to be within the group of *Homo heidelbergensis*, though other designations such as Archaic Homo sapiens and Homo sapiens rhodesiensis have also been proposed.
- In February 2006 a fossil, the Gawis cranium, was found which might possibly be a species intermediate between *H. erectus* and *H. sapiens* or one of many evolutionary dead ends. The skull from Gawis, Ethiopia, is believed to be 500,000–250,000 years old. Only summary details are known, and no peer reviewed studies have been released by the finding team. Gawis man's facial features suggest its being either an intermediate species or an example of a "Bodo man" female.

H. neanderthalensis

Le Ferrassie Neanderthal skull (cast)

H. neanderthalensis lived from 400,000 years ago. Also proposed as *Homo sapiens neanderthalensis*: there is ongoing debate over whether the "Neanderthal Man" was a separate species, *Homo neanderthalensis*, or a subspecies of *H. sapiens*. While the debate remains unsettled, evidence from sequencing mitochondrial DNA indicates that no significant gene flow occurred between *H. neanderthalensis* and *H. sapiens*, and, therefore, the two were separate species that shared a common ancestor about 660,000 years ago. In 1997, Mark Stoneking stated: "These results [based on mitochondrial DNA extracted from Neanderthal bone] indicate that Neanderthals did not contribute mitochondrial DNA to modern humans… Neanderthals are not our ancestors." Subsequent investigation of a second source of Neanderthal DNA supported these findings. However, the 2010 sequencing of the Neanderthal genome indicates that Neanderthals did indeed interbreed with *H. sapiens circa* 75,000 BC (after *H. sapiens* moved out from Africa, but before they separated into Europe, the Middle East, and Asia). Nearly all modern humans have 1% to 4% of their DNA derived from Neanderthal DNA. (To appreciate how big of a percentage this is, consider that humans and chimps only differ in 1.5% of their DNA.) This 1-4% bit of DNA is only present in non-African humans. However, supporters of the multiregional hypothesis point to recent studies indicating non-African nuclear DNA heritage dating to one Ma, although the reliability of these studies has been questioned. Competition from *Homo sapiens* probably contributed to Neanderthal extinction. They could have coexisted in Europe for as long as 10,000 years.

H. sapiens

H. sapiens (the adjective *sapiens* is Latin for "wise" or "intelligent") have lived from about 250,000 years ago to the present. Between 400,000 years ago and the second interglacial period in the Middle Pleistocene, around 250,000 years ago, the trend in skull expansion and the elaboration of stone tool technologies developed, providing evidence for a transition from *H. erectus* to *H. sapiens*. The direct evidence suggests there was a migration of *H. erectus* out of Africa, then a further speciation of *H. sapiens* from *H. erectus* in Africa. A subsequent migration within and out of Africa eventually replaced the earlier dispersed *H. erectus*. This migration and origin theory is usually referred to as the *recent single origin* or Out of Africa theory. Current evidence does not preclude some multiregional evolution or some admixture of the migrant *H. sapiens* with existing *Homo* populations. This is a hotly debated area of paleoanthropology.

Current research has established that humans are genetically highly homogenous; that is, the DNA of individuals is more alike than usual for most species, which may have resulted from their relatively recent evolution or the possibility of a population bottleneck resulting from cataclysmic natural events such as the Toba catastrophe. Distinctive genetic characteristics have arisen, however, primarily as the result of small groups of people moving into new environmental circumstances. These adapted traits are a very small component of the *Homo sapiens* genome, but include various characteristics such as skin color and nose form, in addition to internal characteristics such as the ability to breathe more efficiently in high altitudes.

H. sapiens idaltu, from Ethiopia, is a possible extinct sub-species who lived from about 160,000 years ago.

H. floresiensis

H. floresiensis, which lived from approximately 100,000 to 12,000 before present, has been nicknamed *hobbit* for its small size, possibly a result of insular dwarfism. *H. floresiensis* is intriguing both for its size and its age, being a concrete example of a recent species of the genus *Homo* that exhibits derived traits not shared with modern humans. In other words, *H. floresiensis* share a common ancestor with modern humans, but split from the modern human lineage and followed a distinct evolutionary path. The main find was a skeleton believed to be a woman of about 30 years of age. Found in 2003 it has been dated to approximately 18,000 years old. The living woman was estimated to be one meter in height, with a brain volume of just 380 cm^3 (considered small for a chimpanzee and less than a third of the *H. sapiens* average of 1400 cm^3).

However, there is an ongoing debate over whether *H. floresiensis* is indeed a separate species. Some scientists presently believe that *H. floresiensis* was a modern *H. sapiens* suffering from pathological dwarfism. This hypothesis is supported in part, because some modern humans who live on Flores, the island where the skeleton was found, are pygmies. This coupled with pathological dwarfism, it is argued, could indeed create a hobbit-like human. The other major attack on *H. floresiensis* is that it was found with tools only associated with *H. sapiens*.

The hypothesis of pathological dwarfism, however, fails to explain additional anatomical features that are unlike those of modern humans (diseased or not) but much like those of ancient members of our genus. Aside from cranial features, these features include the form of bones in the wrist, forearm, shoulder, knees, and feet.

Denisova hominin

In 2008, archeologists working at the site of Denisova Cave in the Altai Mountains of Siberia uncovered a small bone fragment from the fifth finger of a juvenile hominin, dubbed the "X-woman" (referring to the maternal descent of mitochondrial DNA, or the Denisova hominin. Artifacts, including a bracelet, excavated in the cave at the same level were carbon dated to around 40,000 BP. As DNA had survived in the fossil fragment due to the cool climate of the Denisova Cave, a team of scientists from the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany sequenced mtDNA extracted from the fragment.

The analysis indicated that modern humans, Neanderthals, and the Denisova hominin last shared a common ancestor around 1 million years ago. Modern humans are known to have overlapped with Neanderthals in Europe for more than 10,000 years, and the discovery raises the possibility that Neanderthals, modern humans and the Denisovan hominin may have co-existed together.

The DNA analysis further indicated that this new hominin species was the result of an early migration out of Africa, distinct from the later out-of-Africa migrations associated with Neanderthals and modern humans, but also distinct from the earlier African exodus of *Homo erectus*. Professor Chris Stringer, human origins researcher at London's Natural History Museum and one of the leading proponents of the recent single-origin hypothesis, remarked: "This new DNA work provides an entirely new way of looking at the still poorly understood evolution of humans in central and eastern Asia." Pääbo noted that the existence of this distant branch creates a much more complex picture of humankind during the Late Pleistocene.

Comparative table of *Homo* species

Comparative table of *Homo* species

Species	Lived when	Lived where	Adult	Adult	Cranial capacity	Fossil	Discovery / publication

	(Ma)		height	mass	(cm³)	record	of name
H. antecessor	1.2 – 0.8	Spain	1.75 m (5.7 ft)	90 kg (200 lb)	1,000	2 sites	1997
H. cepranensis	0.9 – 0.8?	Italy			1,000	1 skull cap	1994/2003
H. erectus	1.5 – 0.2	Africa, Eurasia (Java, China, India, Caucasus)	1.8 m (5.9 ft)	60 kg (130 lb)	850 (early) – 1,100 (late)	Many	1891/1892
H. ergaster	1.9 – 1.4	Eastern and Southern Africa	1.9 m (6.2 ft)		700–850	Many	1975
H. floresiensis	0.10? – 0.012	Indonesia	1.0 m (3.3 ft)	25 kg (55 lb)	400	7 individuals	2003/2004
H. gautengensis	>2 – 0.6	South Africa	1.0 m (3.3 ft)				2010
H. georgicus	1.8	Georgia			600	4 individuals	1999/2002
H. habilis	2.3 – 1.4	Africa	1.0–1.5 m (3.3–4.9 ft)	33–55 kg (73–120 lb)	510–660	Many	1960/1964
H. heidelbergensis	0.6 – 0.35	Europe, Africa, China	1.8 m (5.9 ft)	60 kg (130 lb)	1,100–1,400	Many	1908
H. neanderthalensis	0.35 – 0.03	Europe, Western Asia	1.6 m (5.2 ft)	55–70 kg (120–150 lb) (heavily built)	1,200–1,900	Many	(1829)/1864

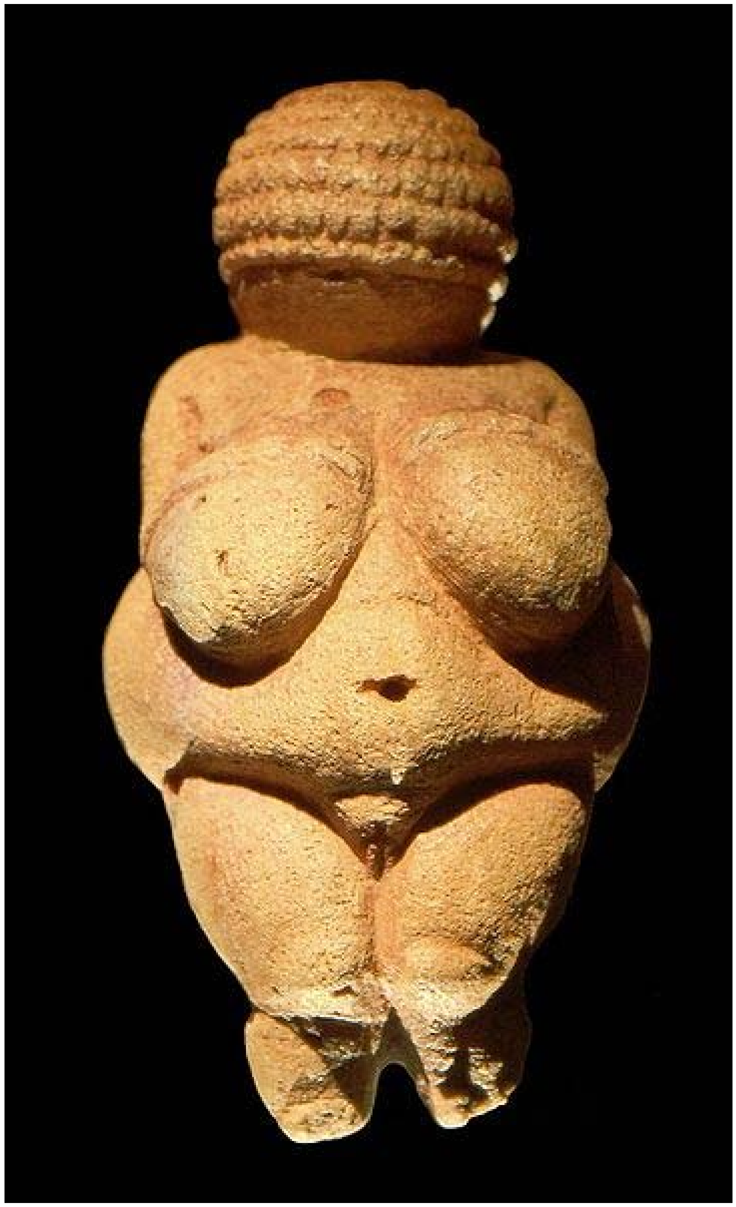
H. rhodesiensis	0.3 – 0.12	Zambia			1,300	Very few	1921
H. rudolfensis	1.9	Kenya				1 skull	1972/1986
H. sapiens idaltu	0.16 – 0.15	Ethiopia			1,450	3 craniums	1997/2003
H. sapiens sapiens (modern humans)	0.2 – present	Worldwide	1.4–1.9 m (4.6–6.2 ft)	50–100 kg (110–220 lb)	1,000–1,850	Still living	—/1758

Use of tools

"A sharp rock", an Oldowan pebble tool, the most basic of human stone tools

Fire, one of the greatest human discoveries

An Acheulean hand axe, the pinnacle of *Homo erectus* stone working



Venus of Willendorf, an example of Paleolithic art

Using tools has been interpreted as a sign of intelligence, and it has been theorized that tool use may have stimulated certain aspects of human evolution—most notably the continued expansion of the human brain. Paleontology has yet to explain the expansion of this organ over millions of years despite being extremely demanding in terms of energy consumption. The brain of a modern human consumes about 20 watts (400 kilocalories per day), which is one fifth of the energy consumption of a human body. Increased tool use would allow hunting for energy-rich meat products, and would enable processing more energy-rich plant products. Researchers have suggested that early hominids were thus under evolutionary pressure to increase their capacity to create and use tools.

Precisely when early humans started to use tools is difficult to determine, because the more primitive these tools are (for example, sharp-edged stones) the more difficult it is to decide whether they are natural objects or human artifacts. There is some evidence that the australopithecines (4 Ma) may have used broken bones as tools, but this is debated.

It should be noted that many species make and use tools, but it is the human species that dominates the areas of making and using more complex tools. The oldest known tools are the "Oldowan stone tools" from Ethiopia. It was discovered that these tools are from 2.5 to 2.6 million years old, which predates the earliest known "Homo" species. There is no known evidence that any "Homo" specimens appeared by 2.5 Ma. A Homo fossil was found near some Oldowan tools, and its age was noted at 2.3 million years old, suggesting that maybe the Homo species did indeed create and use these tools. It is surely possible, but not solid evidence. Bernard Wood noted that "Paranthropus" coexisted with the early Homo species in the area of the "Oldowan Industrial Complex" over roughly the same span of time. Although there is no direct evidence that points to Paranthropus as the tool makers, their anatomy lends to indirect evidence of their capabilities in this area. Most paleoanthropologists agree that the early "Homo" species were indeed responsible for most of the Oldowan tools found. They argue that when most of the Oldowan tools were found in association with human fossils, Homo was always present, but Paranthropus was not.

In 1994, Randall Susman used the anatomy of opposable thumbs as the basis for his argument that both the Homo and Paranthropus species were toolmakers. He compared bones and muscles of human and chimpanzee thumbs, finding that humans have 3 muscles that chimps lack. Humans also have thicker metacarpals with broader heads, making the human hand more successful at precision grasping than the chimpanzee hand. Susman defended that modern anatomy of the human thumb is an evolutionary response to the requirements associated with making and handling tools and that both species were indeed toolmakers.

Stone tools

Stone tools are first attested around 2.6 Ma, when *H. habilis* in Eastern Africa used so-called pebble tools, choppers made out of round pebbles that had been split by simple strikes. This marks the beginning of the Paleolithic, or Old Stone Age; its end is taken to be the end of the last Ice Age, around 10,000 years ago. The Paleolithic is subdivided into the Lower Paleolithic (Early Stone Age, ending around 350,000–300,000 years ago), the Middle Paleolithic (Middle Stone Age, until 50,000–30,000 years ago), and the Upper Paleolithic.

The period from 700,000–300,000 years ago is also known as the Acheulean, when *H. ergaster* (or *erectus*) made large stone hand-axes out of flint and quartzite, at first quite rough (Early Acheulian), later "retouched" by additional, more subtle strikes at the sides of the flakes. After 350,000 BP (Before Present) the more refined so-called Levallois technique was developed. It consisted of a series of consecutive strikes, by which scrapers, slicers ("racloirs"), needles, and flattened needles were made. Finally, after about 50,000 BP, ever more refined and specialized flint tools were made by the Neanderthals and the immigrant Cro-Magnons (knives, blades, skimmers). In this period they also started to make tools out of bone.

Modern humans and the "Great Leap Forward" debate

Until about 50,000–40,000 years ago the use of stone tools seems to have progressed stepwise. Each phase (*H. habilis*, *H. ergaster*, *H. neanderthalensis*) started at a higher level than the previous one, but once that phase started further development was slow. These *Homo* species were culturally conservative, but after 50,000 BC modern human culture started to change at a much greater speed. Jared Diamond, author of *The Third Chimpanzee*, and some anthropologists characterize this as a "Great Leap Forward".

Modern humans started burying their dead, making clothing out of hides, developing sophisticated hunting techniques (such as using trapping pits or driving animals off cliffs), and engaging in cave painting. As human culture advanced, different populations of humans introduced novelty to existing technologies: artifacts such as fish hooks, buttons and bone needles show signs of variation among different populations of humans, something that had not been seen in human cultures prior to 50,000 BP. Typically, *H. neanderthalensis* populations do not vary in their technologies.

Among concrete examples of Modern human behavior, anthropologists include specialization of tools, use of jewellery and images (such as cave drawings), organization of living space, rituals (for example, burials with grave gifts), specialized hunting techniques, exploration of less hospitable geographical areas, and barter trade networks. Debate continues as to whether a "revolution" led to modern humans ("the big bang of human consciousness"), or whether the evolution was more gradual.

Models of human evolution

Today, all humans belong to one, undivided by species barrier, population of *Homo sapiens sapiens*. However, according to the "Out of Africa" model this is not the first species of hominids: the first species of genus *Homo*, *Homo habilis*, evolved in East Africa at least 2 Ma, and members of this species populated different parts of Africa in a relatively short time. *Homo erectus* evolved more than 1.8 Ma, and by 1.5 Ma had spread throughout the Old World.

Anthropologists have been divided as to whether current human population evolved as one interconnected population (as postulated by the Multiregional Evolution hypothesis), or evolved only in East Africa, speciated, and then migrating out of Africa and replaced human populations in Eurasia (called the "Out of Africa" Model or the "Complete Replacement" Model).

Multiregional model

Multiregional evolution, a *model to account for the pattern of human evolution*, was proposed by Milford H. Wolpoff in 1988. Multiregional evolution holds that human evolution from the beginning of the Pleistocene 2.5 million years BP to the present day has been within a single, continuous human species, evolving worldwide to modern *Homo sapiens*.

According to the multiregional hypothesis, fossil and genomic data are evidence for worldwide human evolution and contradict the recent speciation postulated by the Recent African origin hypothesis. The fossil evidence was insufficient for Richard Leakey to resolve this debate. Studies of haplogroups in Y-chromosomal DNA and mitochondrial DNA have largely supported a recent African origin. Evidence from autosomal DNA also supports the Recent African origin. However the presence of archaic admixture in modern humans remains a possibility and has been suggested by some studies.

Out of Africa

According to the Out of Africa model, developed by Chris Stringer and Peter Andrews, modern *H. sapiens* evolved in Africa 200,000 years ago. *Homo sapiens* began migrating from Africa between 70,000 – 50,000 years ago and eventually replaced existing hominid species in Europe and Asia. Out of Africa has gained support from research using mitochondrial DNA (mtDNA). After analysing genealogy trees constructed using 133 types of mtDNA, researchers concluded that all were descended from a woman from Africa, dubbed Mitochondrial Eve. Out of Africa is also supported by the fact that mitochondrial genetic diversity is highest among African populations.

There are differing theories on whether there was a single exodus or several. A multiple dispersal model involves the Southern Dispersal theory, which has gained support in recent years from genetic, linguistic and archaeological evidence. In this theory, there was a coastal dispersal of modern humans from the Horn of Africa around 70,000 years ago. This group helped to populate Southeast Asia and Oceania, explaining the discovery of early human sites in these areas much earlier than those in the Levant. A second wave of humans dispersed across the Sinai peninsula into Asia, resulting in the bulk of human population for Eurasia. This second group possessed a more sophisticated tool technology and was less dependent on coastal food sources than the original group. Much of the evidence for the first group's expansion would have been destroyed by the rising sea levels at the end of the Holocene era. The multiple dispersal model is contradicted by studies indicating that the populations of Eurasia and the populations of Southeast Asia and Oceania are all descended from the same mitochondrial DNA lineages, which support a single migration out of Africa that gave rise to all non-African populations.

The broad study of African genetic diversity headed by Dr. Sarah Tishkoff found the San people to express the greatest genetic diversity among the 113 distinct populations sampled, making them one of 14 "ancestral population clusters". The research also located the origin of modern human migration in south-western Africa, near the coastal border of Namibia and Angola.

According to the Toba catastrophe theory to which some anthropologists and archeologists subscribe, the supereruption of Lake Toba on Sumatra island in Indonesia roughly 70,000 years ago had global consequences, killing most humans then alive and creating a population bottleneck that affected the genetic inheritance of all humans today.

Recent and current human evolution

Natural selection is being observed in contemporary human populations, with recent findings demonstrating the population which is at risk of the severe debilitating disease kuru has significant over-representation of an immune variant of the prion protein gene G127V versus non-immune alleles. Scientists postulate one of the reasons for the rapid selection of this genetic variant is the lethality of the disease in non-immune persons. Other reported evolutionary trends in other populations include a lengthening of the reproductive period, reduction in cholesterol levels, blood glucose and blood pressure.

In their 2009 book *The 10,000 Year Explosion*, Gregory Cochran and Henry Harpending argue that human evolution has accelerated since and as a result of the development of agriculture and civilisation since some 50,000 years ago, and that there are consequently substantial genetic differences between different current human populations.

Genetics

Human evolutionary genetics studies how one human genome differs from the other, the evolutionary past that gave rise to it, and its current effects. Differences between genomes have anthropological, medical and forensic implications and applications. Genetic data can provide important insight into human evolution.

Notable human evolution researchers

- Robert Broom, a Scottish physician and palaeontologist whose work on South Africa led to the discovery and description of the Paranthropus genus of hominins, and of "Mrs. Ples"
- James Burnett, Lord Monboddo, a British judge most famous today as a founder of modern comparative historical linguistics
- Raymond Dart, an Australian anatomist and palaeoanthropologist, whose work at Taung, in South Africa, led to the discovery of *Australopithecus africanus*
- Charles Darwin, a British naturalist who documented considerable evidence that species originate through evolutionary change
- Richard Dawkins, a British ethologist, evolutionary biologist who has promoted a gene-centered view of evolution
- J. B. S. Haldane, a British geneticist and evolutionary biologist

- William D. Hamilton, a British Evolutionary Biologist who expounded a rigorous genetic basis for kin selection, and on the evolution of HIV and other human diseases.
- Sir Alister Hardy, a British zoologist, who first hypothesised the aquatic ape theory of human evolution
- Henry McHenry, an American anthropologist who specializes in studies of human evolution, the origins of bipedality, and paleoanthropology
- Jeffrey Laitman, an American anatomist and physical anthropologist whose work has explored the evolution of the vocal tract and speech
- Louis Leakey, an African archaeologist and naturalist whose work was important in establishing human evolutionary development in Africa
- Mary Leakey, a British archaeologist and anthropologist whose discoveries in Africa include the Laetoli footprints
- Richard Leakey, an African paleontologist and archaeologist, son of Louis and Mary Leakey
- Svante Pääbo, a Swedish biologist specializing in evolutionary genetics
- David Pilbeam, a paleoanthropologist, researcher and writer on a range of topics involving human and primate evolution.
- Jeffrey H. Schwartz, an American physical anthropologist and professor of biological anthropology
- Chris Stringer, anthropologist, leading proponent of the recent single origin hypothesis
- Alan Templeton, geneticist and statistician, proponent of the multiregional hypothesis
- Philip V. Tobias, a South African palaeoanthropologist is one of the world's leading authorities on the evolution of humankind
- Erik Trinkaus, a prominent American paleoanthropologist and expert on Neanderthal biology and human evolution
- Alfred Russel Wallace, a British naturalist, sometimes called the "father of biogeography", who independently from Charles Darwin proposed the principles of evolution of animal species
- Milford H. Wolpoff, an American paleoanthropologist who is the leading proponent of the multiregional evolution hypothesis.

Chapter- 7

History of the World

History of the world (or "World History") is generally taken to encompass the history of the human race (*Homo sapiens*), from the earliest times to the present, in all places on earth. By convention, it generally excludes non-human *natural history* and *geological history*, except insofar as the natural world substantially affects human lives. It encompasses the study of written records, from ancient times forward, plus additional knowledge gained from other sources, such as *archaeology*. Ancient recorded history begins with the invention, independently at several sites on Earth, of writing, which created the infrastructure for lasting, accurately transmitted memories and thus for the diffusion and growth of knowledge. However, the roots of civilization reach back to the period before writing — humanity's prehistory.

Human prehistory begins in the Paleolithic, or Early Stone Age. During the Neolithic (New Stone Age) Agricultural Revolution between 8,500 and 7,000 BCE in the Fertile Crescent, humans began the systematic husbandry of plants and animals — agriculture. It spread to neighboring regions, and also developed independently elsewhere, until most humans lived as farmers in permanent settlements. The relative security and increased productivity provided by farming allowed these communities to expand. They grew over time into increasingly larger units in parallel with the evolution of ever more efficient means of transport.

Surplus food made possible an increasing division of labor, the rise of a leisured upper class, and the development of cities and thus of civilization. The growing complexity of human societies necessitated systems of accounting, which led to writing.

Civilizations developed on the banks of life-sustaining bodies of fresh water (lakes and rivers). By 3,000 BCE they had arisen in the Middle East's Mesopotamia (the "land between the Rivers" Euphrates and Tigris), on the banks of Egypt's River Nile, and in the Indus River valley. Similar civilizations are believed also to have arisen at this time along the great rivers of China, but the archaeological evidence for extensive urban construction is less distinct.

The history of the Old World (Europe in particular) is commonly divided into Antiquity, up to 476 CE; the Middle Ages, from the 5th through the 15th centuries, including the early European Renaissance; the Early Modern period, from the 15th century to the late

18th, including the Age of Enlightenment; and the Modern period, from the Industrial Revolution to the present.

In Europe, the fall of the Western Roman Empire (476 CE) is commonly taken as signaling the end of antiquity and the beginning of the Middle Ages, during which (around the year 1300) the European Renaissance emerges. In the mid-15th century, Johannes Gutenberg's invention of modern printing, employing movable type, revolutionized communication, helping end the Middle Ages and usher in modern times and the Scientific Revolution. By the 18th century, the accumulation of knowledge and technology, especially in Europe, had reached a critical mass that brought about the Industrial Revolution.

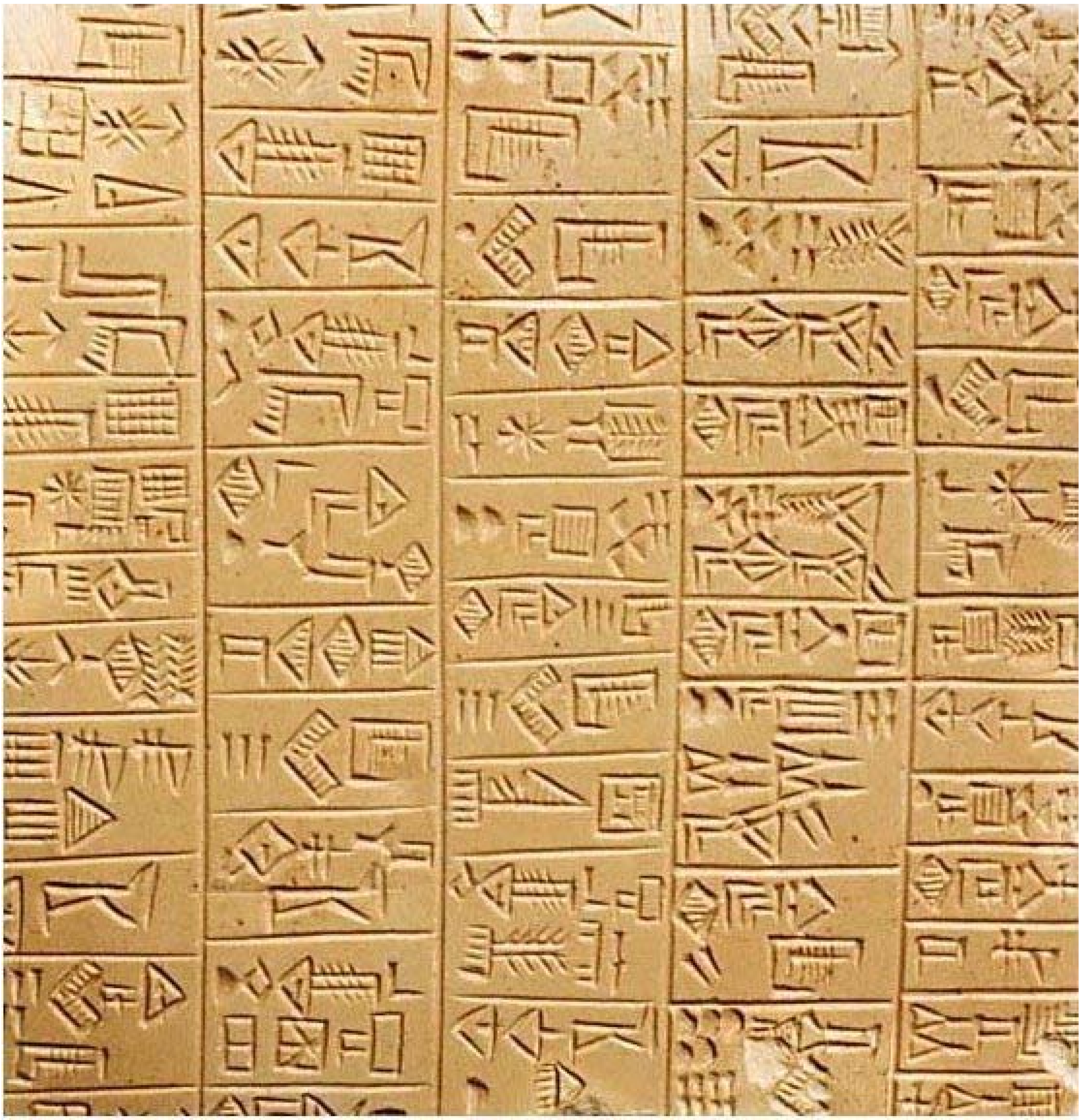
In other parts of the world, in the ancient Near East, ancient China, ancient India and elsewhere, the historic timeline unfolded differently, but by the 18th century, due to extensive world trade and colonization, the histories of the many human civilizations had in large measure converged. In the last quarter-millennium, the growth of knowledge, technology, commerce, and of the potential destructiveness of war has accelerated, creating the opportunities and perils that now confront the many human communities that inhabit the planet.

Prehistory

Homo sapiens first arose on the Earth between 400,000 and 250,000 years ago during the Palaeolithic period. This occurred after a long period of evolution. Ancestors of humans, such as *Homo erectus*, had been using simple tools for many millennia, but as time progressed, tools became far more refined and complex. At some point, humans had begun using fire for heat and for cooking. Humans also developed language in the Palaeolithic, as well as a conceptual repertoire that included systematic burial of the dead and adornment of the living. During this period, all humans lived as hunter-gatherers, who were generally nomadic.

Modern humans spread rapidly from Africa and the frost-free zones of Europe and Asia. The rapid expansion of humankind to North America and Oceania took place at the climax of the most recent Ice Age, when temperate regions of today were extremely inhospitable. Yet, humans had colonised nearly all the ice-free parts of the globe by the end of the Ice Age, some 12,000 years ago.

The Agricultural Revolution, beginning about 10,000 BCE, saw the development of agriculture. Farming permitted far denser populations, which in time organised into states. Agriculture also created food surpluses that could support people not directly engaged in food production. The development of agriculture permitted the creation of the first cities. These were centres of trade, manufacture and political power with nearly no agricultural production of their own. Cities established a symbiosis with their surrounding countrysides, absorbing agricultural products and providing, in return, manufactures and varying degrees of military control and protection.



Cuneiform script, the earliest known writing system

The development of cities equated, both etymologically and in fact, with the rise of civilization itself. Before the age of cities (about 40,000 BC), there is evidence of people living in man-made shelter huts in northern Punjab and central Asia (Bactria). By 7,000 BC, there is evidence of people growing barley in this area and raising sheep and goats. Around this time, people begin living in mud-brick dwellings in villages, some of which are still in existence. Early cities arose in the first Sumerian civilization, in lower Mesopotamia (3500 BCE), followed by Egyptian civilization along the Nile (3300 BCE) and Harappan civilization in the Indus Valley (3300 BCE). Elaborate cities grew up, with high levels of social and economic complexity. Each of these civilizations was so different from the others that they almost certainly originated independently. It was at this time, and due to the needs of cities, that writing and extensive trade were introduced.

This period also saw the origins of complex religion. Religious belief in this period commonly consisted in the worship of a Mother Goddess, a Sky Father, and of the Sun and Moon as deities. Shrines developed, which over time evolved into temple establishments, complete with a complex hierarchy of priests and priestesses and other functionaries. Typical of the Neolithic was a tendency to worship anthropomorphic deities. Some of the earliest surviving written religious scriptures are the *Pyramid Texts*, produced by the Egyptians, the oldest of which date to between 2400 and 2300 BCE. Some archeologists suggest, based on ongoing excavations of a temple complex at Göbekli Tepe ("Potbelly Hill") in southern Turkey, dating from ca. 11,500 years ago, that religion predated the Agricultural Revolution rather than following in its wake, as had generally been assumed.

Ancient history

Origin of civilization

Ancient Egyptians built the Great Pyramids of Giza

The Bronze Age forms part of a three-age system. In this system, in some areas of the world, the Bronze Age follows the Stone Age. During this era the most fertile areas of the world saw city states and the first civilizations develop. These were concentrated on four fertile river valleys: The Tigris and Euphrates in Mesopotamia, the Nile in Egypt, the Indus in South Asia, and the Yangtze and Yellow River in China.

Mesopotamia saw the rise of the Sumerian city states. It was in these cities that the earliest known form of writing, cuneiform script, appeared ca. 3000 BCE. Cuneiform writing began as a system of pictographs. Over time, the pictorial representations became simplified and more abstract. Cuneiforms were written on clay tablets, on which symbols were drawn with a blunt reed for a stylus. Writing made the administration of a large state far easier. This era also saw new military technologies, such as chariots, that allowed armies to move faster.

These developments led to the development of empires. The first empire, controlling a large territory and many cities, developed in Egypt that formed with the unification of Lower and Upper Egypt c. 3100 BCE. Over the next millennia the other river valleys would also see monarchical empires rise to power. In the 24th century BCE, the Akkadian Empire arose in Mesopotamia and in China the Xia Dynasty arose c. 2200 BCE.

Over the next millennia civilizations would develop across the world. Trade would increasingly become a source of power as states with access to important resources or controlling important trade routes would rise to dominance. In c.2,500 BCE the Kingdom of Kerma developed in Sudan south of Egypt. In modern Turkey the Hittites controlled a large empire and by 1600 BCE, Mycenaean Greece began to develop. In India this era was the Vedic period, which laid the foundations of Hinduism and other cultural aspects of early Indian society, and ended in the 500s BCE. From around 550 BCE, many independent kingdoms and republics known as the Mahajanapadas were established across the country. In the Americas, civilizations such as the Maya, Zapotec, Moche, and Nazca emerged in Mesoamerica and Peru at the end of the 1st millennium BCE.

Ancient empires

Religion and philosophy

Angkor Wat temple, Cambodia, early 12th century

Beginning in the sixth century BCE a set of transformative religious and philosophical ideas developed. During this century Chinese Confucianism, Indian Buddhism and Jainism, Persian Zoroastrianism, Ancient Egyptian Monotheism, and Jewish Monotheism all developed. In the fifth century Socrates and Plato would lay the foundations of Ancient Greek philosophy.

In the east, three schools of thought were to dominate Chinese thinking until the modern day. These were Taoism, Legalism and Confucianism. The Confucian tradition, which would attain dominance, looked for political morality not to the force of law but to the power and example of tradition. Confucianism would later spread into the Korean peninsula and toward Japan.

In the west, the Greek philosophical tradition, represented by Socrates, Plato, and Aristotle, was diffused throughout Europe and the Middle East in the 4th century BCE by the conquests of Alexander III of Macedon, more commonly known as Alexander the Great.

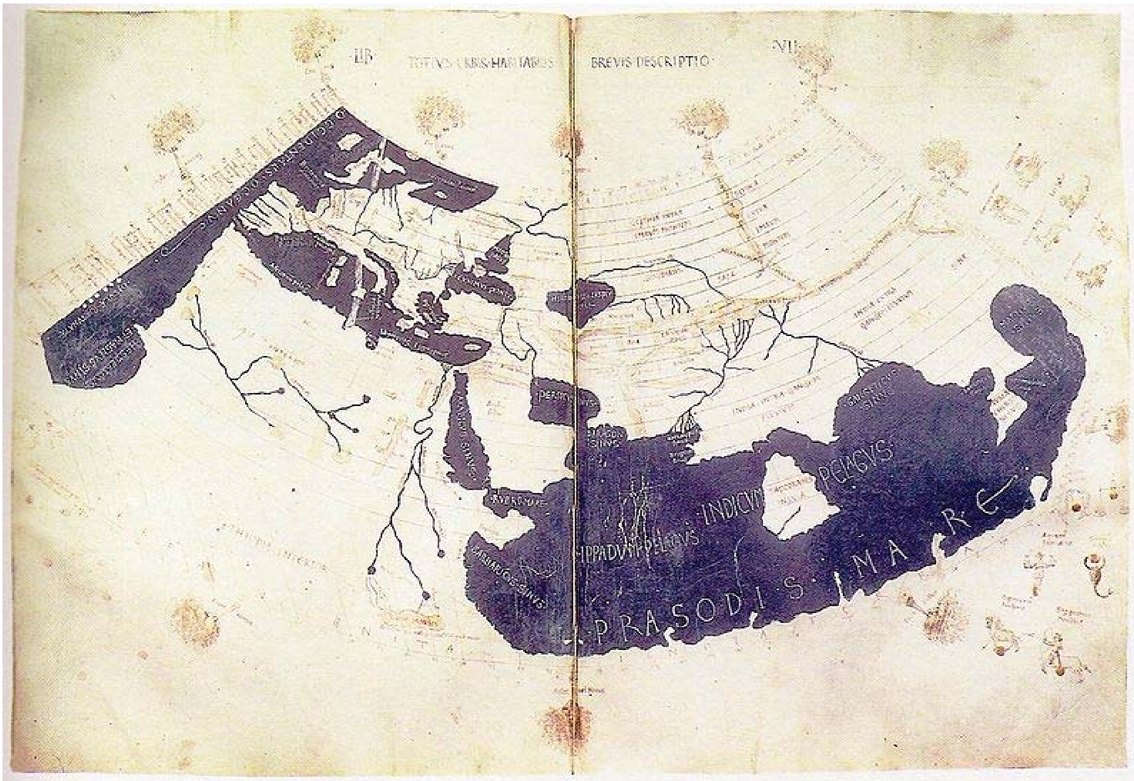
Regional empires

The Parthenon epitomizes the sophisticated culture of the ancient Greeks

The millennia from 500 BCE to 500 CE saw a series of empires of unprecedented size develop. Well-trained professional armies, unifying ideologies, and advanced bureaucracies created the possibility for emperors to rule over domains, whose population could attain numbers upwards of tens of millions of subjects.

This period in the history of the world was marked by slow but steady technological advances, with important developments such as the stirrup and moldboard plow arriving every few centuries. There were, however, in some regions, periods of rapid technological progress. Most important, perhaps, was the Mediterranean area during the Hellenistic period, when hundreds of technologies were invented. Such periods were followed by periods of technological decay, as during the Roman Empire's decline and fall and the ensuing early medieval period.

The great empires depended on military annexation of territory and on the formation of defended settlements to become agricultural centres. The relative peace that the empires brought encouraged international trade, most notably the massive trade routes in the Mediterranean that had been developed by the time of the Hellenistic Age, and the Silk Road.



Ptolemy's world map, reconstituted from his *Geographia* (ca. 150)

The empires faced common problems associated with maintaining huge armies and supporting a central bureaucracy. These costs fell most heavily on the peasantry, while land-owning magnates were increasingly able to evade centralised control and its costs. The pressure of barbarians on the frontiers hastened the process of internal dissolution. China's Han Empire fell into civil war in 220 CE, while its Roman counterpart became increasingly decentralised and divided about the same time.

In the west, the Greeks established a civilization that is the foundational culture of modern western civilization. Some centuries later, in the 3rd century BCE, the Romans began expanding their territory through conquest and colonization. By the reign of Emperor Augustus (late 1st century BCE), Rome controlled all the lands surrounding the Mediterranean. By the reign of Emperor Trajan (early 2nd century CE), Rome controlled much of the land from England to Mesopotamia.

In the third century BCE, most of South Asia was united into the Maurya Empire by Chandragupta Maurya and flourished under Ashoka the Great. From the third century CE, the Gupta dynasty oversaw the period referred to as ancient India's Golden Age. Empires in Southern India included those of the Chalukyas,the Rashtrakutas, the Hoysalas, the Cholas and the Vijayanagara Empire. Science, engineering, art, literature, astronomy, and philosophy flourished under the patronage of these kings.

Declines and falls

The great empires of Eurasia were all located on temperate coastal plains. From the Central Asian steppes, horse-based nomads dominated a large part of the continent. The development of the stirrup, and the breeding of horses strong enough to carry a fully armed archer, made the nomads a constant threat to the more settled civilizations.

The gradual break-up of the Roman Empire, spanning several centuries after the 2nd century CE, coincided with the spread of Christianity westward from the Middle East. The Western Roman Empire fell under the domination of Germanic tribes in the 5th century, and these polities gradually developed into a number of warring states, all associated in one way or another with the Roman Catholic Church. The remaining part of the Roman Empire, in the eastern Mediterranean, would henceforth be the Byzantine Empire. Centuries later, a limited unity would be restored to western Europe through the establishment of the Holy Roman Empire in 962, comprising a number of states in what is now Germany, Austria, Switzerland, Belgium, Italy, and parts of France.

In China, dynasties would similarly rise and fall. After the fall of the Eastern Han Dynasty and the demise of the Three Kingdoms, Nomadic tribes from the north began to invade in the 4th century CE, eventually conquering areas of Northern China and setting up many small kingdoms. The Sui Dynasty reunified China in 581, and under the succeeding Tang Dynasty (618-907) China entered a second golden age. The Tang Dynasty also splintered, however, and after half a century of turmoil the Northern Song Dynasty reunified China in 982. Yet pressure from nomadic empires to the north became increasingly urgent. North China was lost to the Jurchens in 1141, and the Mongol Empire conquered all of China in 1279, as well as almost all of Eurasia's landmass, missing only central and western Europe, and most of Southeast Asia and Japan.

In these times, northern India was ruled by the Guptas. In southern India, three prominent Dravidian kingdoms emerged: Cheras, Cholas and Pandyas. The ensuing stability contributed to heralding in the golden age of Hindu culture in the 4th and 5th centuries CE.

Machu Picchu, "the Lost City of the Incas"—the most recognizable symbol of Inca civilization

Also at this time in Central America, vast societies began to be built, the most notable being the Maya and Aztecs of Mesoamerica. As the mother culture of the Olmecs gradually declined, the great Mayan city-states slowly rose in number and prominence, and Maya culture spread throughout Yucatán and surrounding areas. The later empire of the Aztecs was built on neighboring cultures and was influenced by conquered peoples such as the Toltecs.

In South America, the 14th and 15th centuries saw the rise of the Inca. The Inca Empire of Tawantinsuyu, with its capital at Cusco, spanned the entire Andes Mountain Range. The Inca were prosperous and advanced, known for an excellent road system and unrivaled masonry.

Middle Ages

The Middle Ages are commonly dated from the fall of the Western Roman Empire in the 5th century.

The period corresponds to the Islamic conquests, subsequent Islamic golden age, and commencement and expansion of the Islamic/Arab Slave Trade followed by the Mongol invasions in the Middle East and Central Asia. South Asia saw a series of middle kingdoms of India followed by the establishment of Islamic empires in India. The Chinese Empire saw the succession of the Sui, Tang, Liao, Yuan and Ming Dynasties.

The Black Death was one of the deadliest pandemics in human history. Starting in Asia, the disease reached Mediterranean and western Europe during the late 1340s, and killed tens of millions of Europeans in six years; between a third and a half of the total population.

The Middle Ages witnessed the first sustained urbanization of northern and western Europe. Many modern European states owe their origins to events unfolding in the Middle Ages; present European political boundaries are, in many regards, the result of the military and dynastic achievements during this tumultuous period.

The Middle Ages lasted until the beginning of the Early Modern Period in the 16th century, marked by the rise of nation-states, the division of Western Christianity in the Reformation, the rise of humanism in the Italian Renaissance, and the beginnings of European overseas expansion which allowed for the Columbian Exchange.

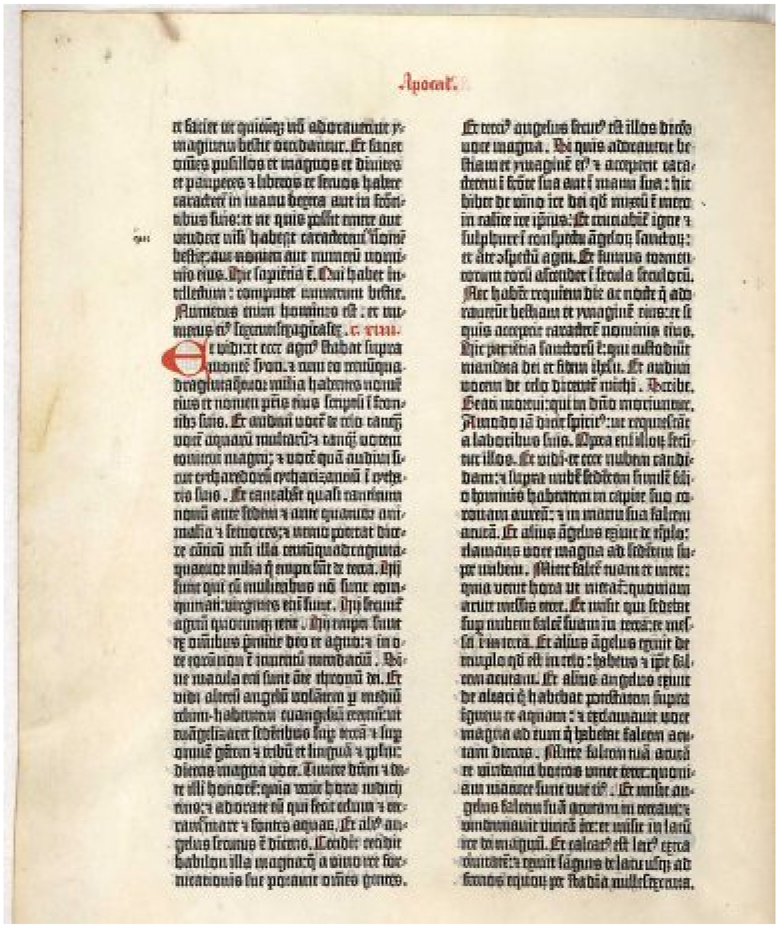
Modern history

Modern history (the "modern period," the "modern era," "modern times") is history of the period following the Middle Ages. "Contemporary history" encompasses historic events that are immediately relevant to the present time; its intentionally loose ambit includes major events such as World War II, but not those whose immediate effects have dissipated.

Early Modern period

"Early modern period" is a term used by historians to refer to the period in Western Europe and its first colonies that spans the centuries between the Middle Ages and the Industrial Revolution. The early modern period is characterized by the rise to importance of science and by increasingly rapid technological progress, secularized civic politics, and the nation-state. Capitalist economies began their rise, initially in northern Italian republics such as Genoa. The early modern period also saw the rise and dominance of the mercantilist economic theory. As such, the early modern period represents the decline and eventual disappearance, in much of the European sphere, of feudalism, serfdom and the power of the Catholic Church. The period includes the Protestant Reformation, the disastrous Thirty Years' War, the European colonization of the Americas, and the peak of European witch-hunting.

Rise of Europe



The movable-type printing press arose in the mid-15th century. Less than 50 years later, nine million books were in print.

Nearly all the agricultural civilizations have been heavily constrained by their environments. Productivity remained low, and climatic changes easily instigated boom and bust cycles that brought about civilizations' rise and fall. By about 1500, however, there was a qualitative change in world history. Technological advance and the wealth generated by trade gradually brought about a widening of possibilities.

Outwardly, Europe's Renaissance, beginning in the 14th century, consisted in the rediscovery of the classical world's scientific contributions, and in the economic and social rise of Europe. But the Renaissance also engendered a culture of inquisitiveness which ultimately led to Humanism, the Scientific Revolution, and finally the great transformation of the Industrial Revolution. The Scientific Revolution in the 17th century, however, had no immediate impact on technology; only in the second half of the 18th century did scientific advances begin to be applied to practical invention.

The advantages that Europe had developed by the mid-18th century were two: an entrepreneurial culture and the wealth generated by the Atlantic trade (including the African slave trade). By the late 16th century, American silver accounted for one-fifth of Spain's total budget. The profits of the slave trade and of West Indian plantations amounted to 5% of the British economy at the time of the Industrial Revolution. While some historians conclude that, in 1750, labour productivity in the most developed regions of China was still on a par with that of Europe's Atlantic economy, other historians like Angus Maddison hold that the per-capita productivity of western Europe had by the late Middle Ages surpassed that of all other regions.

A number of explanations are proffered as to why, from the late Middle Ages on, Europe rose to surpass other civilizations, become the home of the Industrial Revolution, and dominate the world. Max Weber argued that it was due to a Protestant work ethic that encouraged Europeans to work harder and longer than others. Another socioeconomic explanation looks to demographics: Europe, with its celibate clergy, colonial emigration, high-mortality urban centers, periodic famines and outbreaks of the Black Death, continual warfare, and late age of marriage had far more restrained population growth, compared to Asian cultures. A relative shortage of labour meant that surpluses could be invested in labour-saving technological advances such as water-wheels and mills, spinners and looms, steam engines and shipping, rather than fueling population growth.

Many have also argued that Europe's institutions were superior, that property rights and free-market economics were stronger than elsewhere due to an ideal of freedom peculiar to Europe. In recent years, however, scholars such as Kenneth Pomeranz have challenged this view, although the revisionist approach to world history has also met with criticism for systematically "downplaying" European achievements.

Vasco da Gama reached India by sea in 1498

Europe's geography may also have played an important role. The Middle East, India and China are all ringed by mountains but, once past these outer barriers, are relatively flat. By contrast, the Pyrenees, Alps, Apennines, Carpathians and other mountain ranges run through Europe, and the continent is also divided by several seas. This gave Europe some degree of protection from the peril of Central Asian invaders. Before the era of firearms, these nomads were militarily superior to the agricultural states on the periphery of the Eurasian continent and, if they broke out into the plains of northern India or the valleys of China, were all but unstoppable. These invasions were often devastating. The Golden Age of Islam was ended by the Mongol sack of Baghdad in 1258. India and China were subject to periodic invasions, and Russia spent a couple of centuries under the Mongol-

Tatar Yoke. Central and western Europe, logistically more distant from the Central Asian heartland, proved less vulnerable to these threats.

Geography also contributed to important geopolitical differences. For most of their histories, China, India and the Middle East were each unified under a single dominant power that expanded until it reached the surrounding mountains and deserts. In 1600 the Ottoman Empire controlled almost all the Middle East, the Ming Dynasty ruled China, and the Mughal Empire held sway over India. By contrast, Europe was almost always divided into a number of warring states. Pan-European empires, with the notable exception of the Roman Empire, tended to collapse soon after they arose. Another doubtless important geographic factor in the rise of Europe was the Mediterranean Sea, which, for millennia, had functioned as a maritime superhighway fostering the exchange of goods, people, ideas and inventions.

Age of Discovery

Columbus sought India aboard the *Santa Maria* in 1492

In the fourteenth century, the Renaissance began in Europe. Some modern scholars have questioned whether this flowering of art and Humanism was a benefit to science. The era did see an important fusion of Arab and European knowledge. One of the most important developments was the caravel, which combined the Mediterranean lateen sail with European square rigging to create the first vessels that could safely sail the Atlantic Ocean. Along with important developments in navigation, this technology allowed the Italian Christopher Columbus in 1492 to journey across the Atlantic Ocean and bridge the gap between Afro-Eurasia and the Americas.

This had dramatic effects on both continents. The Europeans brought with them viral diseases that American natives had never encountered, and uncertain numbers of natives died in a series of devastating epidemics. The Europeans also had the technological advantage of horses, steel and guns that helped them overpower the Aztec and Incan empires as well as North American cultures.

Gold and resources from the Americas began to be stripped from the land and people and shipped to Europe, while at the same time large numbers of European colonists began to emigrate to the Americas. To meet the great demand for labor in the new colonies, the mass import of Africans as slaves began. Soon much of the Americas had a large racial underclass of slaves. In West Africa, a series of thriving states developed along the coast, becoming prosperous from the exploitation of suffering interior African peoples.

Europe's maritime expansion unsurprisingly — given that continent's geography — was largely the work of its Atlantic states: Portugal, Spain, England, France, and the Netherlands. The Portuguese and Spanish Empires were the predominant conquerors and source of influence, and their union resulted in the Iberian Union, the first global empire, on which the "sun never set". Soon the more northern English, French and Dutch began to dominate the Atlantic. In a series of wars fought in the 17th and 18th centuries, culminating with the Napoleonic Wars, Britain emerged as the new world power.

Meanwhile the voyages of Admiral Zheng He were halted by China's Ming Dynasty (1368–1644), established after the expulsion of the Mongols. A Chinese commercial revolution, sometimes described as "incipient capitalism", was also abortive. The Ming Dynasty would eventually fall to the Manchus, whose Qing Dynasty at first oversaw a period of calm and prosperity but would increasingly fall prey to Western encroachment.

19th century

After Europeans had achieved dominance over the Americas, their imperial appetites turned to the countries of Asia. In the 19th century the European states had a distinct technological advantage over Asian states and peoples. Britain gained control of the Indian subcontinent, Egypt and the Malay Peninsula; the French took Indochina; while the Dutch cemented their control over the Dutch East Indies. In addition, Russia colonised large pre-agricultural areas of Siberia. The British also colonised places inhabited by Neolithic peoples, including Australia, New Zealand and South Africa. large numbers of British colonists emigrated to these colonies. In the late 19th century, the

European powers divided the remaining areas of Africa. Within Europe, economic and military challenges created a system of nation states, and ethno-linguistic groupings began to identify themselves as distinctive nations with aspirations for cultural and political autonomy. This nationalism would become important to peoples across the world in the twentieth century.

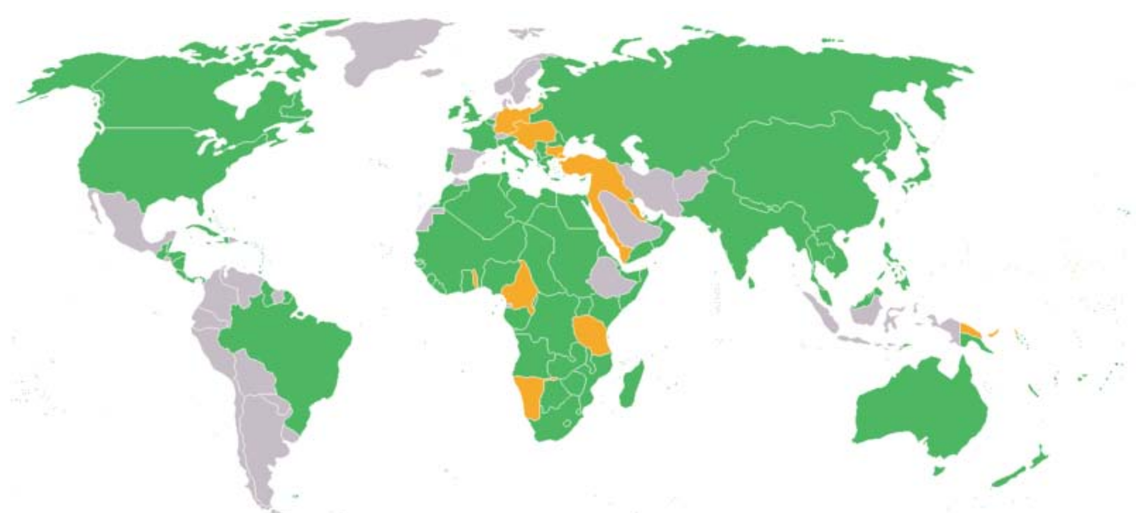
This era in European culture saw the Age of Reason lead to the Scientific Revolution. The Scientific Revolution changed humanity's understanding of the world and happened simultaneously with the Industrial Revolution, a major transformation of the world's economies. The Industrial Revolution began in Great Britain and used new modes of production — the factory, mass production, and mechanisation — to manufacture a wide array of goods faster and using less labour than previously.

The Age of Reason also led to the beginnings of modern democracy in the late-18th century American and French Revolutions. Democracy would grow to have a profound effect on world events and on quality of life.

During the Industrial Revolution, the world economy became reliant on coal as a fuel, as new methods of transport, such as railways and steamships, effectively shrank the world. Meanwhile, industrial pollution and environmental damage, present since the discovery of fire and the beginning of civilization, accelerated drastically.

20th century to present

Early 20th century

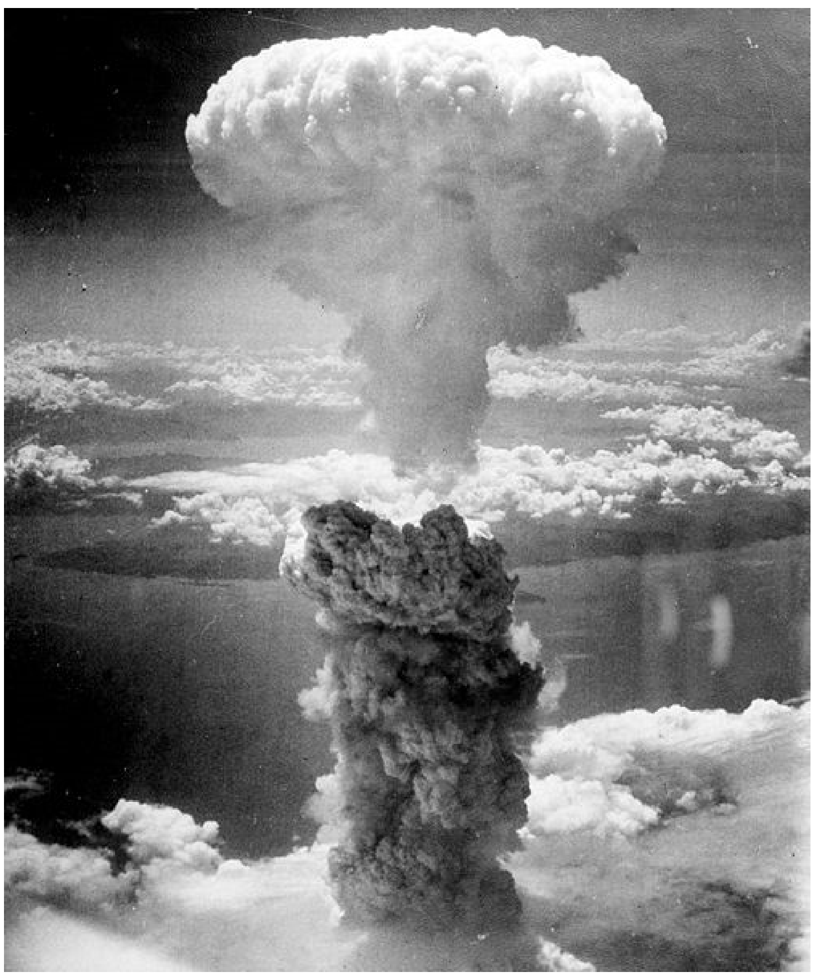


World War I, fought by the Allies *(green)* and Central Powers *(orange)*, ended the German, Austro-Hungarian, Russian and Ottoman Empires.

The 20th century opened with Europe at an apex of wealth and power, and with much of the world under its direct colonial control or its indirect domination. Much of the rest of the world was influenced by heavily Europeanized nations: the United States and Japan. As the century unfolded, however, the global system dominated by rival powers was subjected to severe strains, and ultimately yielded to a more fluid structure of independent nations organized on Western models.

This transformation was catalysed by wars of unparalleled scope and devastation. World War I destroyed many of Europe's empires and monarchies, and weakened Britain and France. In its aftermath, powerful ideologies arose. The Russian Revolution of 1917 created the first communist state, while the 1920s and 1930s saw militaristic fascist dictatorships gain control in Italy, Germany, Spain and elsewhere.

Ongoing national rivalries, exacerbated by the economic turmoil of the Great Depression, helped precipitate World War II. The militaristic dictatorships of Europe and Japan pursued an ultimately doomed course of imperialist expansionism. Their defeat opened the way for the advance of communism into Central Europe, Yugoslavia, Bulgaria, Romania, Albania, China, North Vietnam and North Korea.



Nuclear bombs, dropped on Japan in 1945, ended World War II and opened the Cold War

Following World War II, in 1945, the United Nations was founded in the hope of allaying conflicts among nations and preventing future wars. The war had, however, left two nations, the United States and the Soviet Union, with principal power to guide international affairs. Each was suspicious of the other and feared a global spread of the other's political-economic model. This led to the Cold War, a forty-year stand-off between the United States, the Soviet Union, and their respective allies. With the development of nuclear weapons and the subsequent arms race, all of humanity were put at risk of nuclear war between the two superpowers. Such war being viewed as impractical, proxy wars were instead waged, at the expense of non-nuclear-armed Third World countries.

Late 20th century

The Cold War lasted through to the ninth decade of the twentieth century, when the Soviet Union's communist system began to collapse, unable to compete economically with the United States and western Europe; the Soviets' Central European "satellites" reasserted their national sovereignty, and in 1991 the Soviet Union itself disintegrated. The United States for the time being was left as the "sole remaining superpower".

In the early postwar decades, the African and Asian colonies of the Belgian, British, Dutch, French and other west European empires won their formal independence. These nations faced challenges in the form of neocolonialism, poverty, illiteracy and endemic tropical diseases. Many of the Western and Central European nations gradually formed a political and economic community, the European Union, which subsequently expanded eastward to include former Soviet satellites.

Last Moon landing — Apollo 17 (1972)

The twentieth century saw exponential progress in science and technology, and increased life expectancy and standard of living for much of humanity. As the developed world shifted from a coal-based to a petroleum-based economy, new transport technologies, along with the dawn of the Information Age, led to increased globalization. Space exploration reached throughout the solar system. The structure of DNA, the very template of life, was discovered, and the human genome was sequenced, a major milestone in the understanding of human biology and the treatment of disease. Global literacy rates continued to rise, and the percentage of the world's labor pool needed to produce humankind's food supply continued to drop.

The technologies of sound recordings, motion pictures, and radio and television broadcasting produced a focus on popular culture and entertainment. Television spots sold both commercial products and political candidates. Some worried that the young generation was losing basic literacy skills. Then, in the last decade of this century, a rapid increase took place in the use of personal computers. A global communication network emerged in the Internet. Mass entertainment gave way to individual communication in what has been called a shift from the fourth to a fifth civilization.

The century saw the development of new global threats, such as the spread of nuclear weapons, worldwide epidemics of diseases, global climate change, massive deforestation, overpopulation, and the dwindling of global resources (particularly fossil fuels).

21st century

As the 20th century yielded to the 21st, worldwide demand and competition for resources rose due to growing populations and industrialization, mainly in India, China and Brazil. However, per capita consumption in these countries is far less than in western countries such as the United States of America, Canada, Australia, Britain, etc. This demand is resulting in increased levels of environmental degradation and a growing threat of global warming. This in turn has spurred the development of alternate sources of energy such as solar and other renewable energy varieties, proposals for cleaner fossil-fuel technologies, and consideration of expanded use of nuclear energy.