

Chapter 1.2. Past Environments and Direct Data on Past Evolution

Chapter 1.2 considers direct data on the history of the Earth, including the evolution of life on it, provided by rocks and fossils. Indirect evidence of past evolution, considered in the previous Chapter, are based on inheritance of genetic information from parents to offspring, a process unique to living matter. In contrast, direct evidence of the past are not fundamentally different for living and non-living matter. Of course, it is impossible to understand past evolution of life without some knowledge of the history of non-living matter on Earth.

Section 1.2.1 outlines basic properties of the contemporary Earth which persisted throughout most of its history. A comprehensive coverage of Earth sciences is outside the scope of this book. Thus, only key facts are mentioned, without any attempts to describe in detail how they were discovered.

Section 1.2.2 treats methods of studying past changes in non-living matter on Earth. These methods are based on physics and chemistry. Modern experimental techniques often make it possible determine what and when happened in the past with high precision and confidence.

Section 1.2.3 describes direct evidence of past life and its evolution, provided by fossils. Although fossilization usually is an exception rather than the rule, sedimentary rocks harbor a huge variety of fossils. Being direct messages from the past, fossils often contain information that cannot be obtained by any other means. Still, the fossil record is incomplete, and interpreting fossils and reconstructing past organisms and their evolution can be difficult.

Section 1.2.4 reviews the history of non-living matter on Earth, emphasizing past changes of abiotic conditions experiences by evolving life. Living and non-living matter strongly influenced each other during much of the Earth's history, and some key properties of the contemporary Earth are the result of the activity of life in the past. Overview of data on past life and its environments, presented in the framework of the Geological time scale, is presented at the end of this Section.

Chapter 1.2 complements Chapter 1.1, and together they introduce both kinds of methods that can be used to study evolution of life in the past. Information on past evolution, covered in Chapters 1.3, 1.4, and 1.5, has been obtained using these methods.

Section 1.2.1. Key properties of the Earth

The Earth's orbit is constantly changing, causing quasiperiodic fluctuations in the amount of Solar energy which reaches its surface. The interior of the Earth consists of three parts: an iron-nickel core, the inner part of which is solid and the outer liquid, a silicium- and magnesium-rich solid mantle and a thin crust. An intermediate layer of mantle, the asthenosphere, is capable of slow movements. As the result, lithosphere, the hard outer layer of the Earth consisting of the crust and the upper portion of the mantle, is subdivided into plates that slowly move relative to each other, causing changes in geography and relief. Rocks that constitute the crust can be igneous, sedimentary, and metamorphic, and their transformations into each other can be viewed as a rock cycle.

1.2.1.1. The Earth's orbit

If the Earth and the Sun were dimensionless material points, if no other planets existed, and if there were no relativistic effects, the Earth's orbit would exactly obey Kepler's laws: it would be elliptical with constant parameters, with the Sun located in one of the two foci of the ellipse. However, this is not the case, mostly because the Earth and the Sun are (approximately) small spheres, and not just points. Their diameters are small, relative to the size of the Earth's orbit, but still with non-trivial. This leads to slow changes of the parameters of the Earth's orbit, with important long-term consequences.

Changes of the following three parameters of the Earth's orbit are particularly important: 1) eccentricity, the ratio of the longer over the shorter dimension of the ellipse (at any moment the Sun stays pretty much at one of its current foci), 2) obliquity of the axis of rotation of the Earth around itself, defined as the angle between this axis and the ecliptic, the plane within which the orbit lies, and 3) positions of the Earth at the Equinoxes, which determine, in particular, whether it is summer or winter in the northern hemisphere when the Earth is in the aphelion. Obviously, eccentricity describes how the Earth rotates around the Sun, and the other two parameters describe how the Earth rotates around itself. All the three parameters change in a quasiperiodic way (Fig. 1.2.1.1.a).

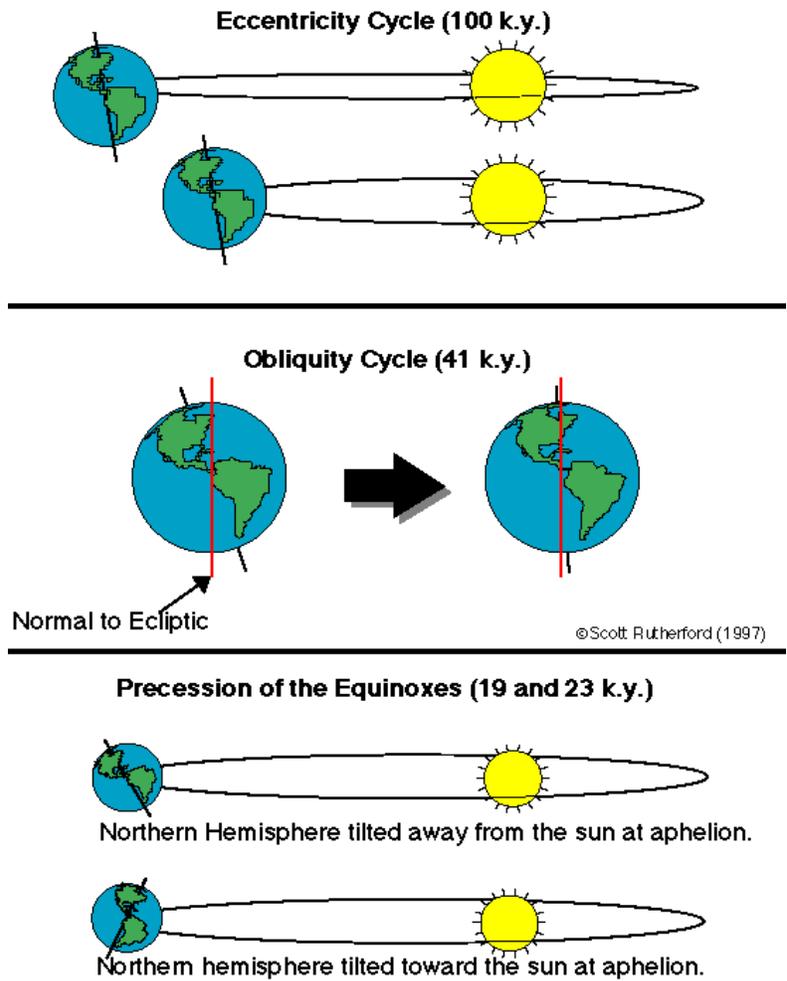


Fig. 1.2.1.1a. Changes in the parameters of Earth's orbit: eccentricity, obliquity, and the positions of the equinoxes. Lengths of time after which the dynamics of a parameter almost repeat themselves are shown in parentheses (*Science* 194, 1121, 1976).

As long as we believe in uniformitarianism (Section 1.1.1.1), slow changes of these three parameters, known as Milankovich's cycles, can be predicted for the future, and reconstructed for the past, by solving Newtonian equations on the motion of two gravitating spheres. Although Newtonian equations are deterministic, we can infer these changes only for up to ~20 Ma in either direction, because attempts to probe deeper into past or future are thwarted by accumulation of errors. Using data on climates in not-too-distant past, for which the parameters of the Earth's orbit can be reconstructed computationally, we can see that past changes in the Earth temperature and the history of glaciations were, to a large extent, due to Milankovich's cycles (Fig. 1.2.1.1b).

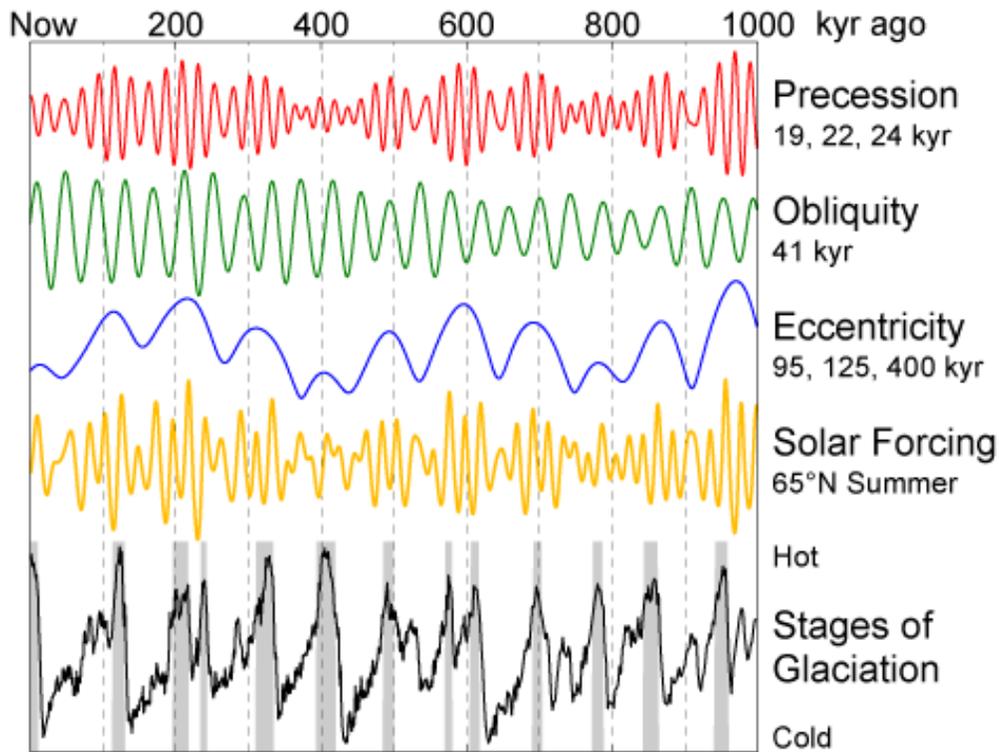


Fig. 1.2.1.1b. The impact of changes of the parameters of the Earth's orbit on the climate.

In addition to these quasiperiodic changes, the Earth's orbit is also subject to an irreversible trend: rotation of the Earth around its axis gradually slows down, due to dissipation of energy caused by tides. The Moon stopped rotating, relative to the Earth, a long time ago, but the more massive Earth remains far from this state. Still, a leap second has to be added occasionally to the Universal time at the end of a year, reflecting slowing down of the Earth's rotation, and the duration of a day grew by ~2 hours in the course of Phanerozoic eon (Section 1.2.4.5). Thus, in late Neoproterozoic era there were ~400 days in a year, because the duration of the year in absolute time did not change noticeably.

1.2.1.2. Structure of the Earth

Although we cannot observe the interior of the Earth directly, indirect data and the hypothetico-deductive method (Section 1.1.1.1) revealed a lot about it. Among these data, propagation of seismic waves is of particular importance (Fig. 1.2.1.2a). Other sources of

information on the interior of the Earth are its magnetic field, gravity, and the emission of heat.

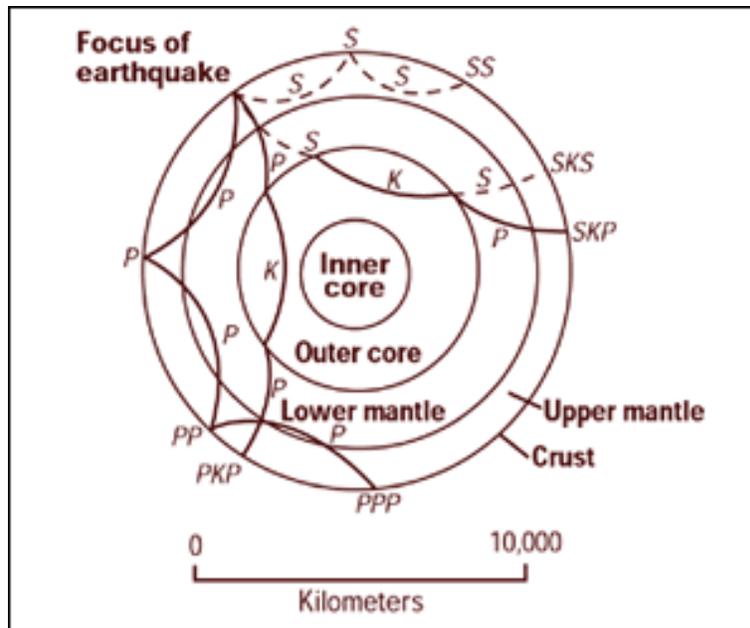


Fig. 1.2.1.2a. Propagation of different kinds of seismic waves. Primary (P) and secondary (S), which are “X-raying” the Earth after each earthquake.

On the basis of this information, there is a general agreement that the Earth consists of the following layers (Fig. 1.2.1.2b):

I. Crust. Outermost solid shell, from 5 (ocean floors) to 80 (under Tibet) km thick, made of silicate rocks, such as basalt (ocean floors) and granite (continents). Crust is separated from mantle by Mohorovicic discontinuity (Moho) where velocity of seismic waves suddenly increases, due to a drastic change in chemical composition of rocks.

II. Mantle. A very thick layer, extending from below the crust to a depth of 2900 km. The mantle is composed of silicate rocks that are richer in iron and magnesium than the overlying crust. The temperature increases from 500°C-900°C at the upper boundary of the mantle to >4,000°C at its lower boundary. Still, the mantle is almost exclusively solid, due to enormous lithostatic pressure. There are 3 distinct layers within the mantle:

1) Uppermost mantle, which extends to the depth of ~100km and is resistant to deformation and relatively cool. Together with the crust, uppermost mantle forms the lithosphere.

2) Asthenosphere (weak sphere), the least rigid layer of the mantle, which extends from the depth of ~100 to 250km.

3) Lower mantle, again a more rigid layer, located below the asthenosphere.

III. Core. The innermost part of the Earth consists primarily of iron and nickel and is separated from mantle by a sharp boundary. The core consists of two layers, liquid outer core, and solid inner core. The innermost part of the core might be composed of very heavy elements, including gold, mercury and uranium. Convection in the outer core gives rise to the Earth magnetic field.

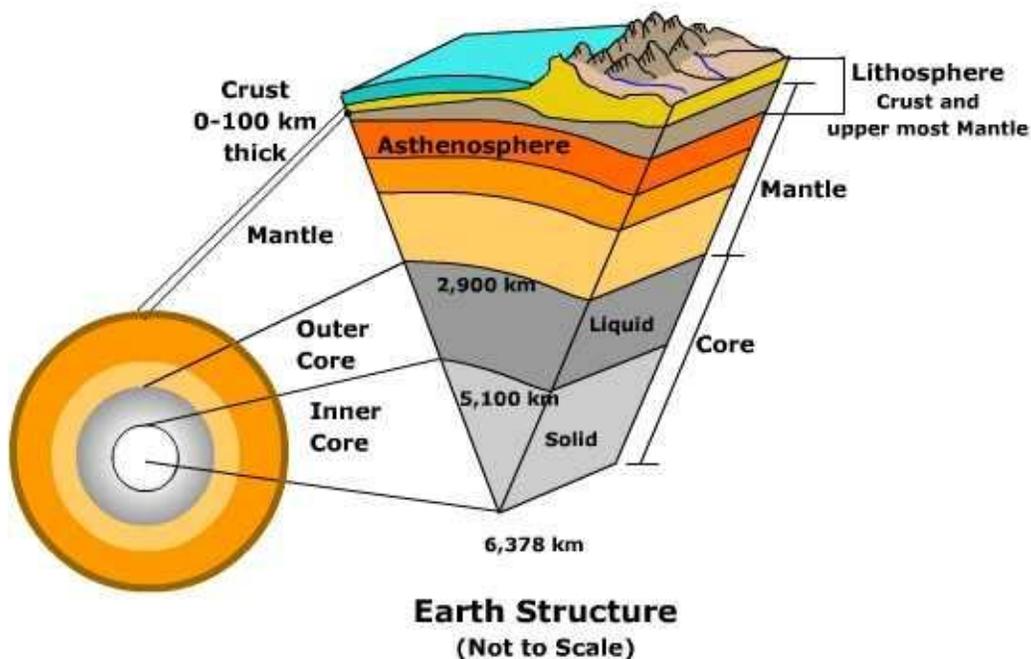


Fig. 1.2.1.2b. Structure of the Earth.

Why do we care about this? Although life inhabits only the upper portion of the Earth crust, processes that occur deep inside the Earth affect life in a variety of ways. The most important of them are plate tectonics and the formation of rocks of which the crust consists. Generation of the heat deep inside the Earth, due to the decay of ^{40}K and, to a lesser extent, of other radioactive isotopes, can be also sensed on its surface (Fig. 1.2.1.2c), although its contribution to the overall thermal balance, and the impact on life, is relatively small.



Fig. 1.2.1.2c. Emission of heat, generated deep inside the Earth, on its surface.

Changes of the Earth's magnetic field, caused by still poorly understood processes in the Earth's core, are also of some importance for biology. Most of the time, geomagnetic poles are approximately aligned to the axis of the Earth's rotation, but occasionally they switch their positions in the course of thousands of years, with the magnetic field becoming much weaker during a reversal (Fig. 1.2.1.2d). These reversals are used to determine the ages of rocks (Section 1.2.2.1).

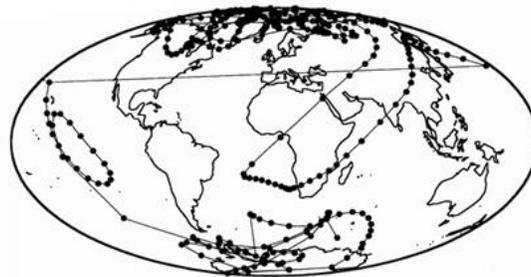
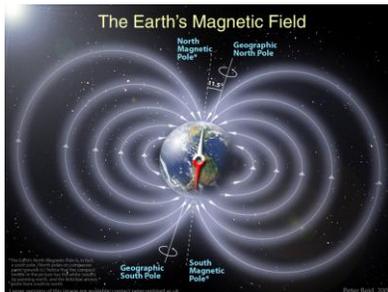


Fig. 1.2.1.2d. The Earth's magnetic field (left) and movements of geomagnetic poles during one geomagnetic polarity reversal even (right), recorded in continuously deposited sediments (*Nature* 389, 712, 1997).

1.2.1.3. Plate tectonics

The lithosphere consists of distinct tectonic plates (Fig. 1.2.1.3a), which ride on the fluid-like (visco-elastic solid) asthenosphere. Plate motions range from a few millimeters to about 15 centimeters per year and have been detected directly. Dissipation of heat from the mantle is the source of energy that drives plate tectonics.

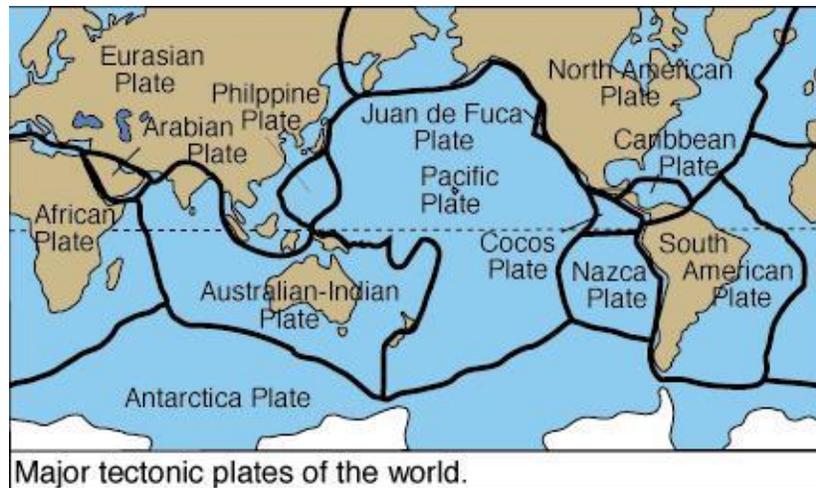


Fig. 1.2.1.3a. 14 major tectonic plates (top). Bridge across the Álfagjá rift valley in southwest Iceland, the boundary of the Eurasian and North American tectonic plates (bottom).

There could be 3 kinds of boundaries between adjacent tectonic plates (Fig. 1.2.1.3b). Transform boundaries occur where the plates slide (grind) past each other. Divergent boundaries occur where two plates slide apart from each other, *e. g.*, at mid-ocean ridges and active zones of rifting, such as the East Africa rift. Convergent boundaries occur where two plates slide towards each other, forming either a subduction zone (if one plate moves underneath the other) or a continental collision (if the two plates contain continental crust). Because of friction and heating, volcanism is almost always closely linked, as in the Andes mountain range and in the Japanese island arc.

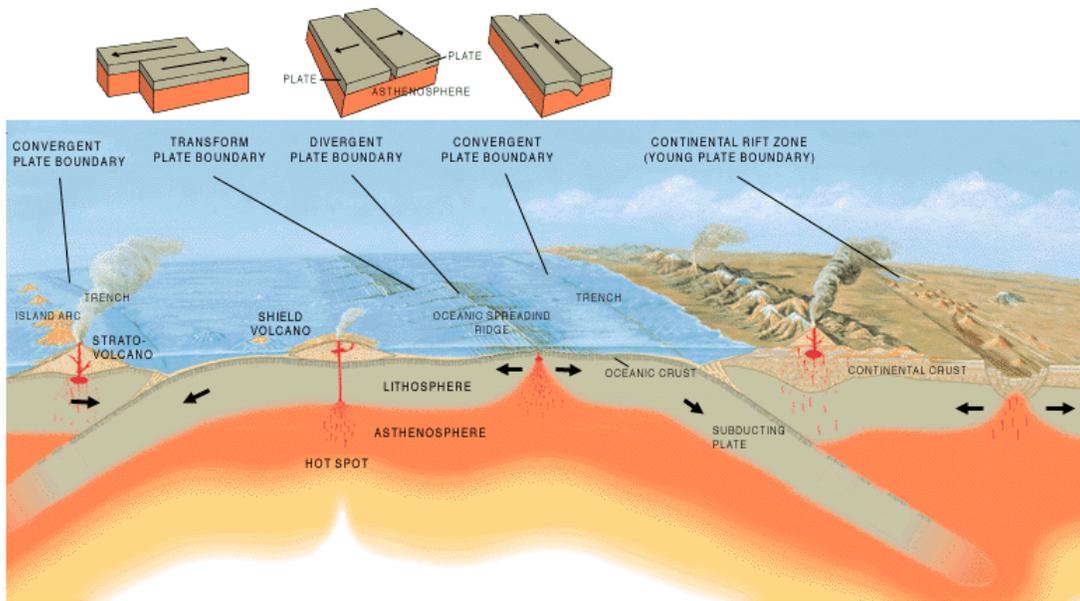


Fig. 1.2.1.3b. Boundaries between plates and processes caused by plate tectonics.

Plate tectonics is responsible for continental drift, and the formation of mountains (orogeny), usually occurring at rate not more than 1mm per year. Another process that affects the face of the earth is erosion. Together, they produce slow but profound changes in geography (Section 1.2.4.2), which strongly affected climate and life on Earth.

1.2.1.4. Types of rocks and the rock cycle

Rocks from which the Earth's crust is composed belong to three distinct classes: igneous, sedimentary, and metamorphic (Fig. 1.2.1.4a). Igneous rocks are rocks solidified from hot silicate melt, magma, or lava. Small crystals in igneous rocks indicate fast cooling, and large crystals indicate slow cooling. Sedimentary rocks originated from layers of deposited particles on the Earth's surface. Continuous deposition of sediments in one location sometimes occurs for many millions of years. Fossils are preserved primarily within sedimentary rocks. Metamorphic rocks are all rocks that changed substantially after their primary formation. Less substantial changes of sedimentary rocks are called diagenesis.



Fig. 1.2.1.4a. Examples of igneous, sedimentary, and metamorphic rocks.

Rocks of these 3 classes can be transformed into each other, in processes that together are known as the Rock Cycle Fig. (1.2.1.4b). Igneous rocks originate from melted material that is transported from the interior of the Earth. Newly formed igneous rocks can penetrate pre-existing rocks, a phenomenon called intrusion. Rocks at the surface decompose/disintegrate by reaction with the atmosphere/hydrosphere to produce solid debris and soluble chemicals that are transported/deposited to form sediments that, upon burial, are converted to sedimentary rocks. Previously formed rocks that are heated and pressurized when buried to shallow to moderate depths (5 to 70 km) become metamorphic rocks. Of course, there is no real periodicity here, and an igneous or a sedimentary rock may persist for billions of years.

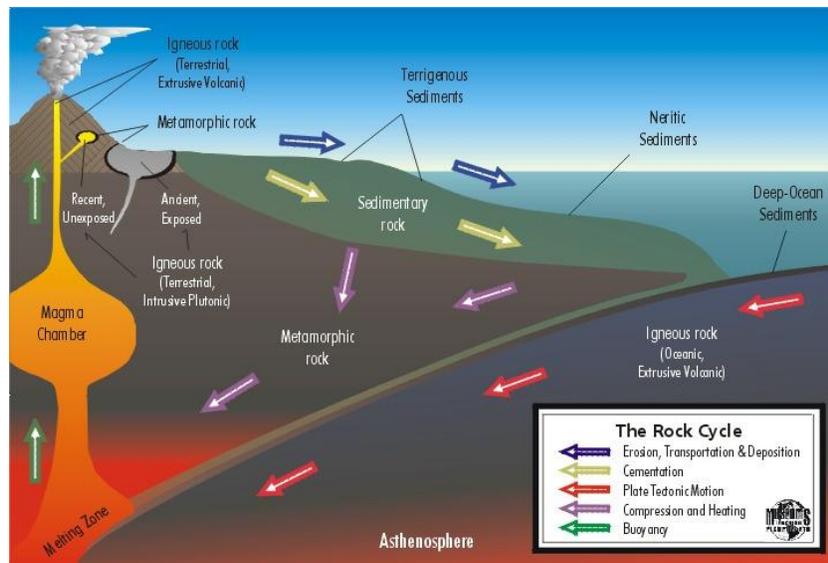


Fig. 1.2.1.4b. The Rock Cycle.

Section 1.2.2. Studying history of the Earth

Non-living matter on the Earth changes relentlessly, mostly slowly but occasionally abruptly, and these changes strongly affected the evolving life. Studies of these changes rely on a wide variety of physical and chemical methods. Both irreversible and repetitive processes can be used to determine the ages of igneous and other rocks. Application of general principles of stratigraphy, together with highly sophisticated analyses, makes it possible to interpret rather complex successions of layers of sedimentary rocks. Data on paleoclimates, paleomagnetism, and on the ranges of ancient forms of life can be used to reconstruct past changes in the locations of continents and in their relief. Chemical features of rocks, including the presence or absence of some compounds and ratios of concentrations of different stable isotopes of several elements, provide information on the parameters of paleoenvironments, such as temperature and composition of the atmosphere. A variety of evidence demonstrates that the Earth experienced a number of global drastic events, some of which profoundly affected life.

1.2.2.1. Geochronology

It is essential to know when past events occurred. Several methods can be used to determine the absolute age of a rock, a fossil, or another dumb object. Although neither of them is universally applicable, together they often provide a reliable estimate of the time of a past event. Let us briefly consider some of these methods.

Radiometric dating. The most important method of determining the age of a dumb object is based on an irreversible process of radioactive decay. For a large number of atoms, radioactive decay is essentially a deterministic process, whose rate, for a particular radioactive isotope, can be measured today with high precision. Because this rate is essentially invariant under physical and chemical condition that can be encountered on the Earth, we can play back the dynamics of radioactive decay (Section 1.1.1.1) and, thus, calculate the age of the crystal from concentrations of parent and daughter isotopes within a sample. Indeed, the dynamics of the number of radioactive atoms in a sample P obeys the following differential equation:

$$\frac{dP}{dt} = -\lambda P \quad (1.2.2.1a)$$

where t is time and the decay constant λ is the probability of an atom decaying during a short unit of time. In contrast to a living being, an atom does not remember its age, and decays with the same probability, regardless of whether it is young or old (Fig. 1.2.2.1a).

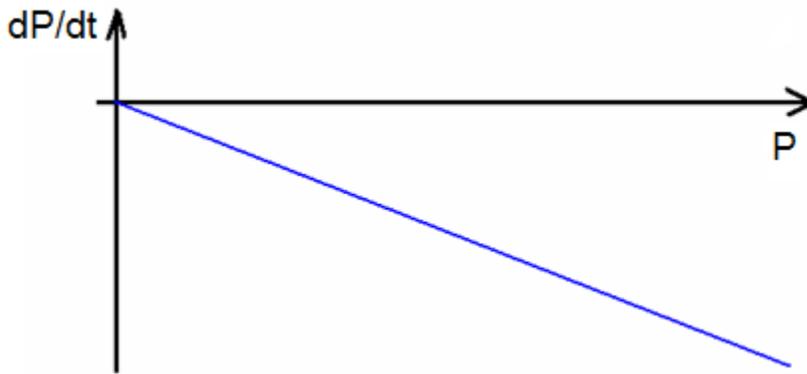


Fig. 1.2.2.1a. Dynamics of decay in a 1-dimensional phase space of the number P of radioactive atoms in a sample. The rate of decline of P , λP , is proportional to it.

Let us solve this equation, *i. e.*, find functions $P(t)$, describing changes of P in time, that obeys it. This exercise prepares us for solving a slightly more difficult differential equation that describes the dynamics of an allele replacement under positive selection (Chapter 2.1). Generally, a solution $x(t)$ of the autonomous differential equation $dx/dt = f(x)$ which satisfies the condition that at some moment of time t_0 the value of the dependent variable x was x_0 can be found by integrating this equation:

$$\int_{x_0}^{x(t)} \frac{dy}{f(y)} = \int_{t_0}^t d\tau = t - t_0 \quad (1.2.2.1b)$$

where y and τ are dummy integration variables. In our case, this general rule reduces to:

$$\int_{P_0}^{P(t)} \frac{dy}{y} = \lambda \int_{t_0}^t d\tau = \lambda(t - t_0) \quad (1.2.2.1c)$$

Thus,

$$\log\left(\frac{P(t)}{P_0}\right) = \lambda(t - t_0) \quad (1.2.2.1d)$$

and

$$P(t) = P_0 e^{-\lambda(t-t_0)} \quad (1.2.2.1e)$$

Formula (1.2.2.1e) provides infinite number of solutions of (1.2.2.1a), each with its own value of P , P_0 , at time t_0 (Fig. 1.2.2.1b). If P_0 is the initial number of atoms of the radioactive isotope at the moment $t = 0$, (1.2.2.1e) is reduced to $P(t) = P_0 e^{-\lambda t}$. We will encounter another case of exponential decay in Chapter 3.1, where divergence of two sequences will be considered quantitatively.

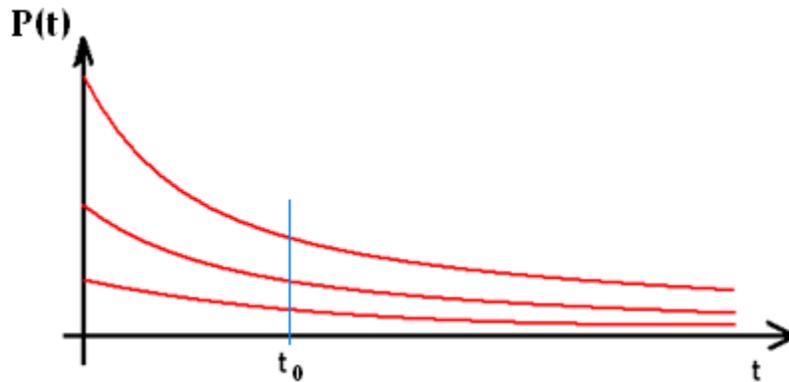


Fig. 1.2.2.1b. A family of functions $P(t)$ describing exponential decline in the number of atoms of a radioactive isotope P , with each function having its own value P_0 at time t_0 .

If the product of decay is stable, and the number of daughter atoms D is known, together with P , we can use this result to measure time, because at every moment of time $P + D = P_0$. Thus,

$$D(t) = P_0(1 - e^{-\lambda t}) \quad (1.2.2.1f)$$

and

$$\frac{D(t)}{P(t)} = \frac{P_0(1 - e^{-\lambda t})}{P_0 e^{-\lambda t}} = e^{\lambda t} - 1 \quad (1.2.2.1g)$$

from which time that elapsed from the moment $t_0 = 0$ when decay started can be recovered:

$$t = \frac{1}{\lambda} \ln\left(1 + \frac{D}{P}\right) \quad (1.2.2.1h)$$

This is the basic equation of radiometric dating. In particular, the time τ that is required for exactly a half of the isotope to decay, known as half-decay period, can be found from (1.2.2.1h), by assuming that $D = P$:

$$\tau = \ln(2) / \lambda \quad (1.2.2.1i)$$

Applying radiometric dating to real rocks requires solving a number of technical problems. In particular, it is essential to be sure that neither parent nor daughter atoms entered or left the sample after its formation. Also, daughter atoms must either be initially absent, or their initial concentration must be somehow estimated.

Many parent-daughter isotope pairs are used in radiometric dating, including $^{238}\text{U}/^{206}\text{Pb}$, $^{235}\text{U}/^{207}\text{Pb}$, $^{40}\text{K}/^{40}\text{Ar}$, $^{87}\text{Rb}/^{87}\text{Sr}$, $^{147}\text{Sm}/^{143}\text{Nd}$, $^{187}\text{Re}/^{187}\text{Os}$, $^{10}\text{Be}/^{10}\text{B}$, and $^{14}\text{C}/^{14}\text{N}$. Parent isotopes in these pairs undergo decay by different mechanisms and have different decay constants. In the case of the first two pairs, there is a number of short-

lived intermediate products between the parent and the daughter isotopes. Let us consider three examples.

Potassium-argon dating is based on the $^{40}\text{K}/^{40}\text{Ar}$ isotope pair. Potassium on the Earth exists in 3 isotopes: ^{39}K (93.2581%), ^{40}K (0.0117%), and ^{41}K (6.7302%). The radioactive isotope ^{40}K decays into ^{40}Ar and ^{40}Ca with a half-decay period of 1.26×10^9 years. As a gas, argon escapes from molten rock and is thus absent within a crystal at the moment of its formation. However, after the rock solidifies, radiogenic ^{40}Ar begins to accumulate in the crystal lattices. Still, the outer, weathered layers of a crystal have to be removed to obtain its accurate age. The most precise way of measuring concentrations of ^{40}K and ^{40}Ar in a sample is to irradiate it with neutrons, converting a known fraction of ^{39}K into ^{39}Ar . The abundances of the radiogenic daughter nuclide ^{40}Ar and of ^{39}Ar (a proxy for the parent nuclide ^{40}K) can be measured in the same sample by mass spectrometry. This approach is known as argon-argon dating, but is just a different technical implementation of potassium-argon dating. Argon-argon dating can be very precise, and can be applied to dating even rather recent events, despite a very long half-decay period of ^{40}K , as well as to dating the most ancient rocks known. As a test, argon-argon method was used to date the eruption of Vesuvius which destroyed Pompeii in 75AD: enough ^{40}Ar to be measured accumulated in solidified lava since the eruption. Of course, the opposite is not true, and isotopes with short half-decay periods cannot be used to date ancient rocks, because after, say, 30 half-decay periods, less than 1 atom out of a billion remains undecayed.

Still, short-lived isotopes can be very useful for dating recent events. In particular, radiocarbon dating is based on $^{14}\text{C}/^{14}\text{N}$ pair, and the half-decay period of ^{14}C is only 5730 years. Radiocarbon dating works differently from potassium-argon dating, because radioactive isotope ^{14}C is present as an approximately constant fraction of all carbon in the atmosphere (two other carbon isotopes, ^{12}C and ^{13}C , are stable), being constantly formed in the upper atmosphere. Thus, the fraction of carbon in a specimen represented by ^{14}C tells us when the specimen died (and, thus, stopped incorporating carbon from atmospheric CO_2 and/or contemporary organic sources), and there is no need to measure the amount of daughter ^{14}N , which would be impossible. Radiocarbon dating can be

applied to organic material younger than ~100 Ky. Obviously, contamination of a sample by more recent material can lead to underestimated dates and must be avoided.

The third example is dating based on spontaneous fission of ^{238}U nuclei into a variety of lighter nuclei. The amount of ^{238}U in a crystal is to be determined chemically. However, instead of attempting to chemically detect all the products of their decay, the decayed ^{238}U atoms are counted, under an optical microscope, by fission tracks left in a crystal by individual decay events (Fig. 1.2.2.1c): a huge amount of energy released in each event is enough to make a sub-macroscopic track.



Fig. 1.2.2.1c. Fission tracks of ^{238}U . Counting these tracks is equivalent to counting the number of decayed nuclei ^{238}U since the crystal was formed, because melting destroys the tracks.

Fossils often are not suitable for radiometric dating, and their age has to be estimated indirectly, from the age of rocks in which they are embedded. In other cases, the age of a fossil can be constrained by the ages of layers of volcanic ash or other rock amenable to radiometric dating present above and below the fossil. Very young fossils can often be dated using ^{14}C .

Other irreversible processes. There are two other irreversible processes that can be used to establish ages of objects: electron capture and racemization. Natural crystals contain "traps" (defects) which can hold electrons excited by radiation. When a crystal is

heated, these trapped electrons are released, emitting a small amount of light (thermoluminescence, Fig. 1.2.2.1d), proportional to the total dose of radiation accumulated by the crystal since the moment when it was last heated or exposed to light. This amount of light can be used to estimate this dose of radiation which in turn contains information on time that lapsed since that moment.

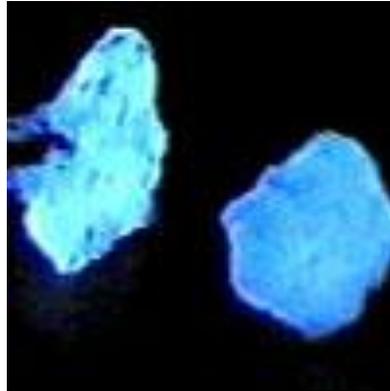


Fig. 1.2.2.1d. Thermoluminescence.

Racemization of amino acids is the process of interconversion of amino acids from one chiral form (L, as only laevo amino acids are present in proteins) to a mixture of L and D (dextro) forms, following protein degradation. At equilibrium there is typically a 1:1 ratio of L to D forms and this mixture is said to be racemic. The known dynamics of increase of the proportion of D-amino acids, as function of time and temperature, can be used for estimating the age of a sample (Fig. 1.2.2.1e).

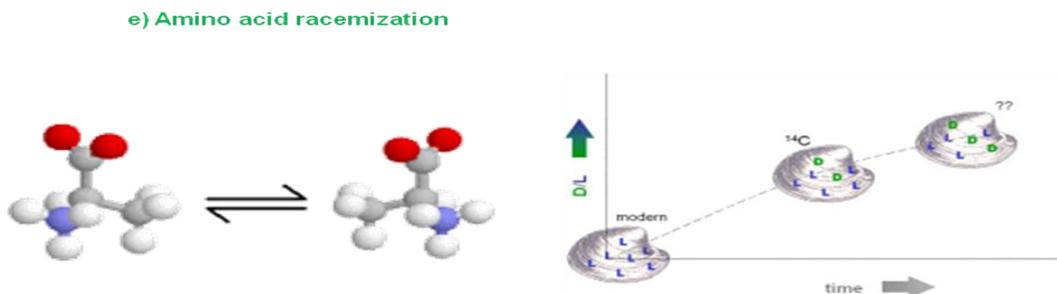


Fig. 1.2.2.1e. Racemization of amino acids is another method of dating, based on an irreversible process.

Repetitive processes. A variety of repetitive processes can be used for dating dumb objects. Layers that correspond to years can be counted in tree trunks, ice cores, and sediments. Of course, in the first case one can look only up to ~5000 years backward, the age of oldest trees, and in the second only up to ~1My, the maximal age of ice in Antarctica. In contrast, sediments sometimes accumulate continuously for many millions of years. Records of temperature changes, provided by stable isotope ratios (Section 1.2.2.4) can be related to such changes calculated on the basis of Milankovich cycles, and their correspondence can be used to date sediments younger than 20My. Relating data on paleomagnetism to the know record of reversals of the Earth's magnetic poles (compiled from numerous measurements around the Globe) makes it possible to determine whether a rock was formed during a period of normal or reverse polarity (Fig. 1.2.2.1f).

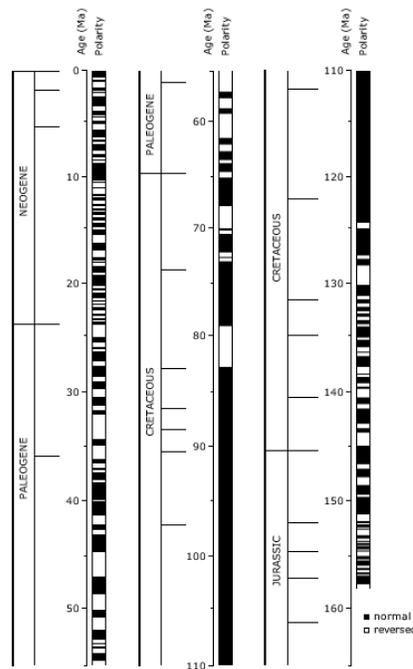


Fig. 1.2.2.1f. Reversals of polarity of the Earth's magnetic field.

1.2.2.2. Stratigraphy

Sedimentary rocks, and to some extent metamorphic rocks derived from them, are of particular importance for studies of past life, because they harbor fossils and contain information about past environments. Sedimentary rocks consist of many layers of deposited material, often corresponding to years. More or less continuous deposition in

one location can go on for hundreds of millions of years, although usually uninterrupted successions of layers cover much shorter periods (Fig. 1.2.2.2a). Sometimes, minute details are preserved in sediments (Fig. 1.2.2.2b).



Fig. 1.2.2.2a. Flat-lying sedimentary rocks from Kaibab Limestone (Permian) at top to Bright Angel Shale (Cambrian) at base of section.



Fig. 1.2.2.2b. Ripple Marks on sandstone of the Triassic Chinle Formation.

Interpretation of sedimentary rocks is the subject of stratigraphy, which is based on several simple principles:

- 1) Superposition: in a sequence of layered rocks, layers are arranged in a time sequence, with the oldest on the bottom and the youngest on the top, unless later processes disturbed this arrangement.
- 2) Original horizontality: layers of sediments are originally deposited horizontally.

3) Lateral continuity: layers of sediments initially extend laterally in all directions. Thus, rocks that are otherwise similar, but are now separated by a valley or other erosional feature, can be assumed to be originally continuous.

4) Biotic succession: different strata contain particular types of fossilized flora and fauna, and these fossil forms succeed each other in a specific and predictable order that can be identified over wide distances.

Interpretation of sedimentary rocks can be complicated by patterns that appear due to processes different from continuous deposition of sediments. Such patterns include:

- 1) hiatuses, caused by ceasing of deposition for some period of time,
- 2) nonconformities, boundaries between igneous and sedimentary rocks,
- 3) angular unconformities, boundaries between two superpositional sedimentary rocks with layers deposited at different angles (Fig. 1.2.2.2c).



Fig. 1.2.2.2c. An angular unconformity between Tertiary sedimentary rocks (tilted beds) and loosely consolidated Quaternary conglomerates and sandstones (horizontal beds) near Pacific Beach on the Olympic Peninsula, Washington.

Nevertheless, it is often possible to interpret, using a variety of methods, even very complex successions of layers of sedimentary rocks, such as those in the Hadar Formation in Ethiopia (Fig. 1.2.2.2d), where a number of crucial hominid fossils were found (Chapter 1.4). Such interpretations are essential for the analysis of fossils.

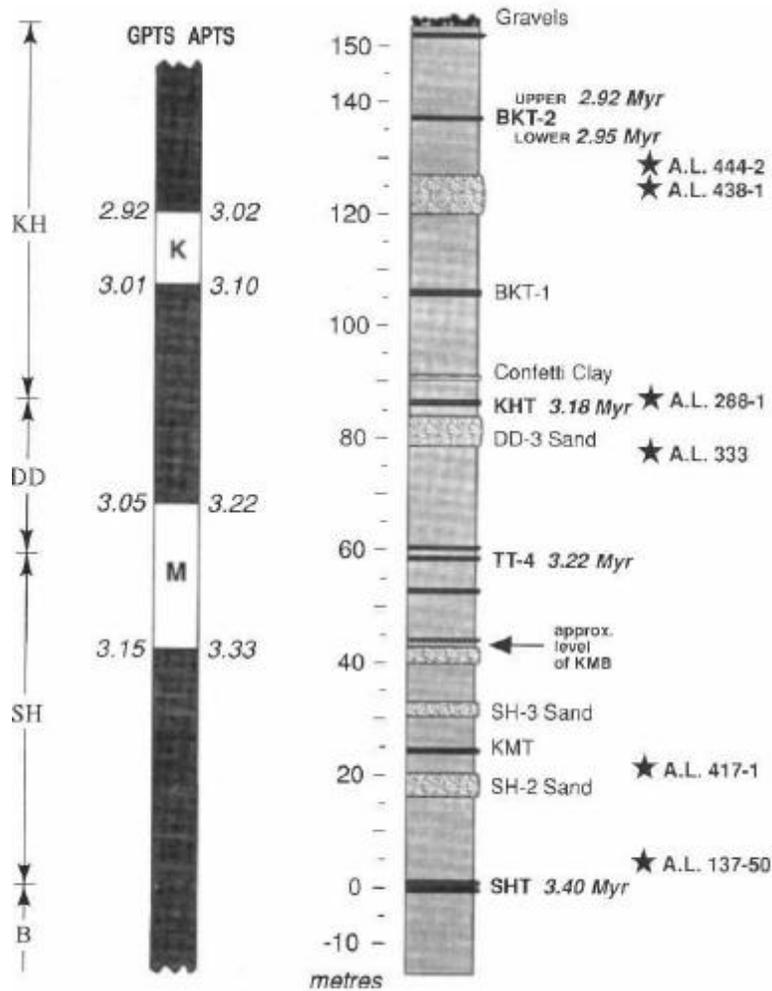


Fig. 1.2.2.2d. Composite stratigraphy of the Hadar Formation, with the positions of hominid fossils indicated by asterisks.

1.2.2.3. Continental drift

The face of the Earth is constantly changing, and this process is ongoing from the very beginning of existence of the Earth's crust. Even major features, including locations of oceans and continents and their relief, are not invariant. Several methods are used to study paleogeography:

- 1) Tectonic reconstructions, based on knowledge of movements of tectonic plates,
- 2) Data on paleomagnetism, which reveal orientation of forming rocks relative to the contemporary positions of magnetic poles,
- 3) Data on local paleoclimates, which may provide information on the geographical latitude of a particular location in the past,

4) Biological correlations (provincialism), which reveal ancient continuity of ranges of species.

This last method illustrates, together with the principle of biotic succession, the utility of fossils for all Earth Sciences. The basic idea is very simple: lands populated by the same (or tightly related) organisms at some moment in the past must have been in close proximity to each other at that moment (Fig. 1.2.2.3a).

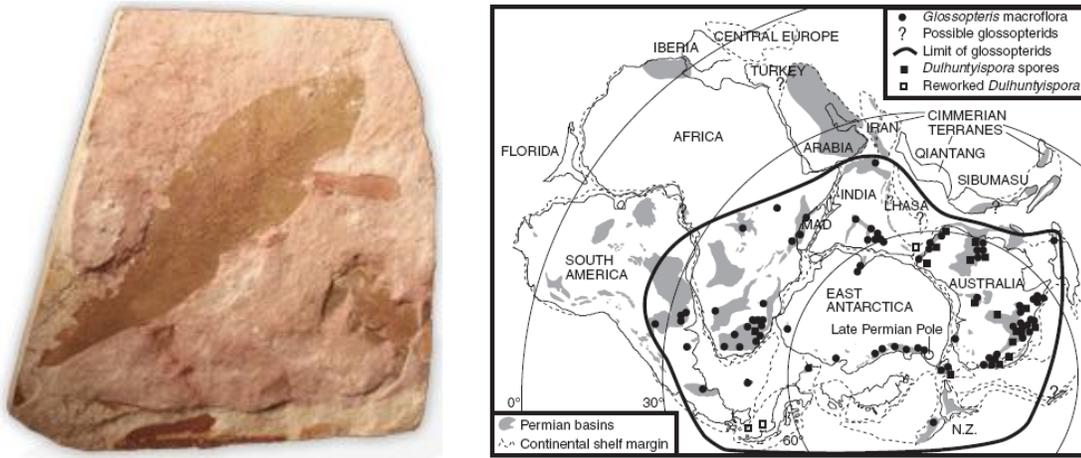


Fig. 1.2.2.3a. An example of provincialism provided by *Glossopteris*, a seed fern that was widely distributed over the southern land masses from ~300 to ~200 mya. Fossils of *Glossopteris* are currently found in Antarctica, Australia, India, and parts of Africa and South America, indicating that all these lands were parts of the Southern supercontinent Gondwana (Section 1.2.4.2).

1.2.2.4. Paleoenvironments

During most of the history of the Earth, environments experienced by past life were very different from the current environments. Parameters of past environments can be inferred from data of several kinds. Let us consider two of them.

Stable isotope ratios. Relative abundances of different stable isotopes of an element in a deposited material depends on the conditions at the time of deposition. For example, there are three stable isotopes of oxygen, ^{16}O , ^{17}O , and ^{18}O , and the ratio of molar concentrations of ^{16}O and ^{18}O is conventionally characterized by

$$\delta^{18}\text{O} = \left(\frac{{}^{18}\text{O}:{}^{16}\text{O}_{\text{sample}} - {}^{18}\text{O}:{}^{16}\text{O}_{\text{standard}}}{{}^{18}\text{O}:{}^{16}\text{O}_{\text{standard}}} \right) \times 1000 \quad (1.2.2.4a)$$

where $^{18}\text{O}:^{16}\text{O}_{\text{standard}}$ is 2005.20. $\delta^{18}\text{O}$ depends on a number of circumstances, mostly because a lighter H_2^{16}O evaporates at a higher rate. Thus, $\delta^{18}\text{O}$ in a sedimentary rock or ancient ice is affected by the overall volume of ice on the Earth at the corresponding time (the ice is enriched by ^{18}O , so that when a lot of ice is present, the remaining water is enriched with ^{16}O) and by the temperature at a site of the deposition.

Several other pairs of stable isotopes, including $^{11}\text{B}/^{10}\text{B}$, $^{13}\text{C}/^{12}\text{C}$, $^{87}\text{Sr}/^{86}\text{Sr}$, and $^{98}\text{Mo}/^{95}\text{Mo}$ are also important for paleoenvironmental reconstructions, and the ratios of their concentrations are characterized analogously to those of the isotopes of oxygen. Sometimes, it is not absolutely clear what caused a rapid change in the ratio of stable isotopes of an element, but the very fact of such a change implies that the global environment changed drastically (Fig. 1.2.2.4a).

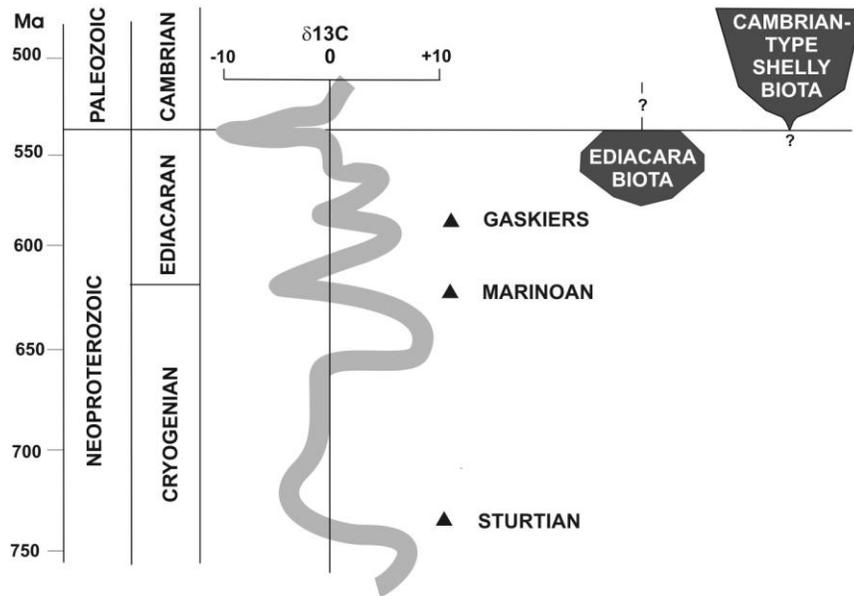


Fig. 1.2.2.4a. A drastic temporary reduction in $\delta^{13}\text{C}$ at the boundary between Neoproterozoic and Paleozoic eons (and Ediacaran and Cambrian periods) indicates rapid changes of the global environment. These changes probably triggered a replacement of Ediacaran biota with Cambrian biota (Section 1.3.2.1). Black triangles mark the three major Neoproterozoic glaciations (*Ann. Rev. Earth and Planetary Sci.* 33, 421, 2005).

In addition to $^{13}\text{C}/^{12}\text{C}$ ratio, $^{34}\text{S}/^{32}\text{S}$ and $^{98}\text{Mo}/^{95}\text{Mo}$ ratios also changed drastically at around the beginning of the Cambrian period. There was a dramatic "signal" in the ratio of two stable molybdenum isotopes, very soon after that time (Fig. 1.2.2.4b). Intense upwelling of H_2S -rich deep ocean water, can explain this signal. If this explanation is correct, there may be a short period of global drop in oxygen concentration in all Ocean water, at the very Ediacaran/Cambrian boundary, followed by oxygenation of the Ocean, including deep water.

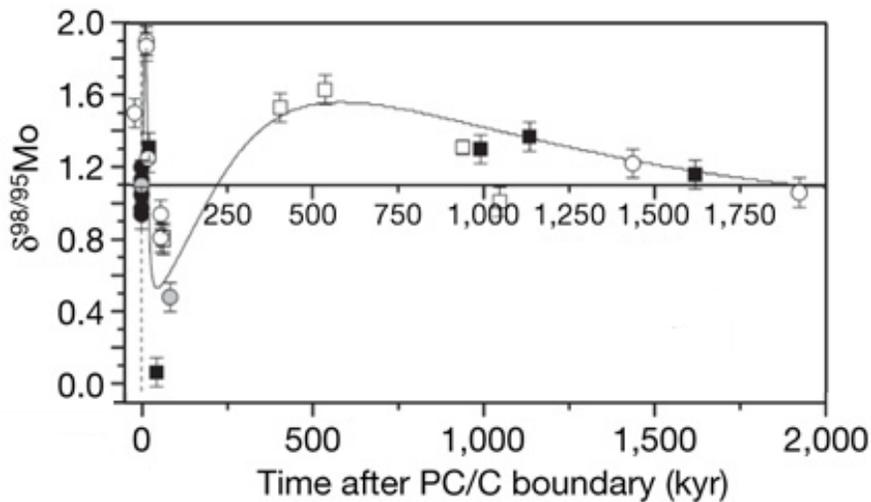


Fig. 1.2.2.4b. Molybdenum isotope signature of black shales possess a transient signal immediately after the Ediacaran/Cambrian (PC/C) boundary (*Nature* 453, 767, 2008).

Chemical composition of sediments. Data of this kind can shed light on paleoenvironments in a variety of ways. Increased abundances of several elements indicated high activity of volcanoes at the time of deposition. The presence of some chemical compounds carries information on the makeup of the atmosphere. For example, mineral uraninite, becomes oxidized if exposed to free oxygen, and is present in deposits formed only before 2,000 mya, when the concentration of O_2 in the atmosphere was below 1%. Many chemical biogenic chemical compounds are bioindicators, revealing the presence of organisms that produced them (Section 1.2.3.2).

1.2.2.5. Drastic events

All studies of the past are based on uniformitarianism. However, invariance of the laws of nature does not necessarily imply that the Earth did not experience global, drastic changes. In fact, such changes occurred many times, sometimes making disastrous impacts on the global biota. Indeed, while new forms of life cannot appear too fast, a mass extinction caused by a global disaster can be instant at the geological time scale. There was a number of mass extinctions (Section 1.3.2.8) and all, or almost all, of them were probably triggered by drastic, global changes of the environment.

Different episodes of drastic changes of the environment were apparently due to different causes. One of such causes are impacts of sizeable extraterrestrial bodies, which happened truly instantaneously (Fig. 1.2.2.5a). An impact leaves a variety of footprints that can be recognized billions of years later.



Fig. 1.2.2.5a. A likely picture of an extraterrestrial impact striking the ocean.

Impact craters. There are 46 known craters with diameters above 20km left by extraterrestrial impacts on the Earth's crust. Most of these craters are buried under the sediments, and their total number is definitely much higher, as the surface of the Moon suggests. The oldest and largest of them is Vredefort crater in South Africa, 2,023 Ma old and 250-300km in diameter (Fig. 1.2.2.5b). Careful studies of such craters reveal a lot of macroscopic features suggestive of their origin. A salient example of an impact crater is Chicxulub (Fig. 1.2.2.5c), 180 km in diameter and 65 Ma old, produced by a meteorite with ~ 10 km. An energy released on this impact was equivalent to ~100,000 gigatons of TNT, a million times more than the most powerful explosive device ever detonated. The Chicxulub impact was a likely cause of the KT extinction (Section 1.3.2.8).

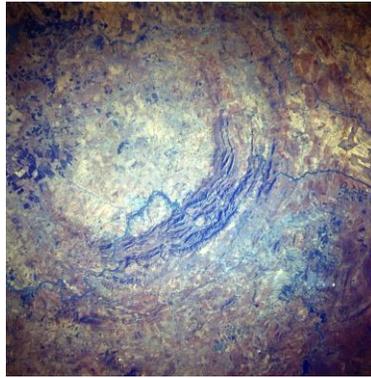


Fig. 1.2.2.5b. Vredefort impact crater (*Processes on the Early Earth* 405, 33, 2006).

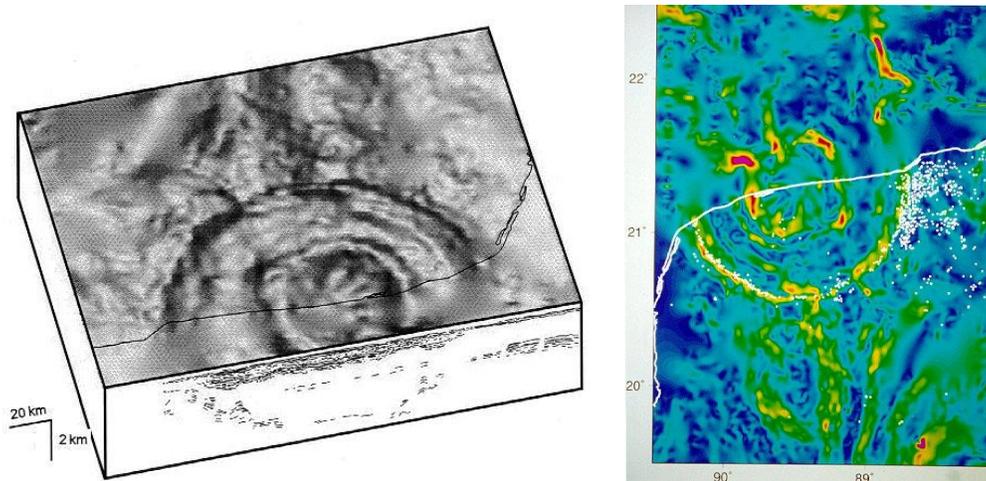


Fig. 1.2.2.5c. Chicxulub is an impact crater buried partially underneath the Yucatán Peninsula and partially under the sea. Reconstruction of the crater (left) and its structure revealed by slight changes in the intensity of gravitation on the surface, which reflect different degrees of compression of underlying rocks (right).

Iridium anomalies. Concentration of iridium in the Earth's crust is very low, presumably because this element easily mixes with iron and mostly sunk to the core of the Earth. However, there are many thin layers of sediments, often found globally, in which iridium is present at anomalously high concentrations (up to 100 times normal). As iridium is known to be much more abundant in asteroids and comet bodies than in the Earth crust, an iridium anomaly is evidence of a large impact, that has spread iridium-rich dust around the world. There are several well-documented iridium anomalies, including

the one at the Cretaceous-Tertiary (K-T) boundary, almost certainly due to the Chicxulub impact. Another substantial iridium anomaly occurred at the Triassic-Jurassic boundary 200Mya (Fig. 1.2.2.5d).

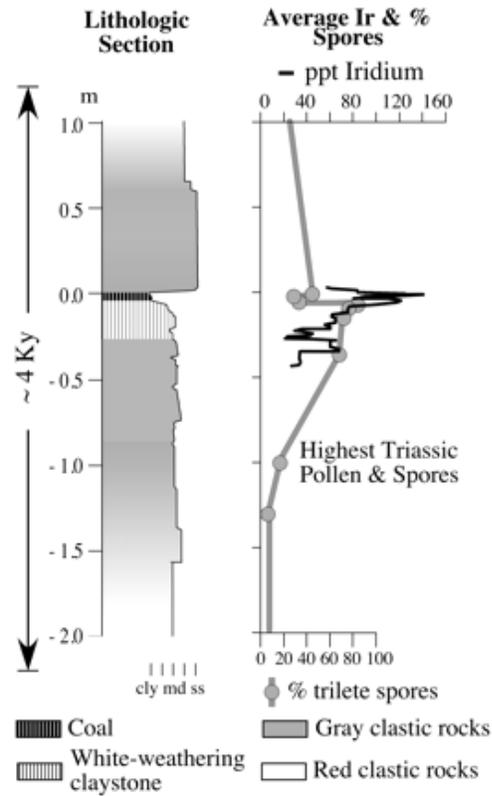


Fig. 1.2.2.5d. Iridium anomaly at the Triassic-Jurassic boundary. A spike in iridium concentration in coincides with a thin layer of coal and is observed in deposits that accumulated in the course of only a few hundred years (*Science* 296, 1305, 2002).

Shocked quartz and microtektites of impact-melt glass. An impact also leaves a lot of microscopic traces. Crystals of quartz that were subject to an impact shock retain a recognizable structure. Droplets of rocks melted by an impact solidify to form impact-melted glass, or microtektites (Fig. 1.2.2.5e).

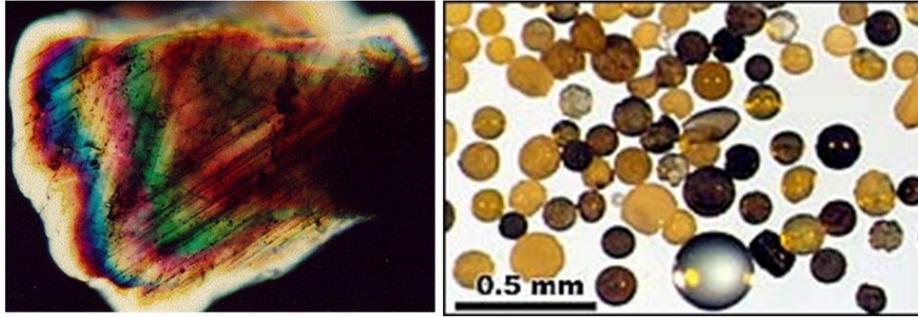


Fig. 1.2.2.5e. Shocked quartz (left) is a form of quartz that has a peculiar microscopic structure. Under intense pressure (but limited temperature), the crystalline structure of quartz is deformed along planes inside the crystal. These planes, called shock lamellae, show up as lines under a microscope. Microtektites (right) are submillimeter-size dark-colored, rounded silicate glass particles, often aerodynamically shaped.

Tsunami deposits. A large enough meteorite that lands in the Ocean must cause gigantic tsunamis, up to 1 km in height, spreading in all directions. Deposits left by such tsunamis are known starting from very early times (Fig. 1.2.2.5f). Traces of tsunamis caused by the Chicxulub impact demonstrate that Cuba was hit especially hard.

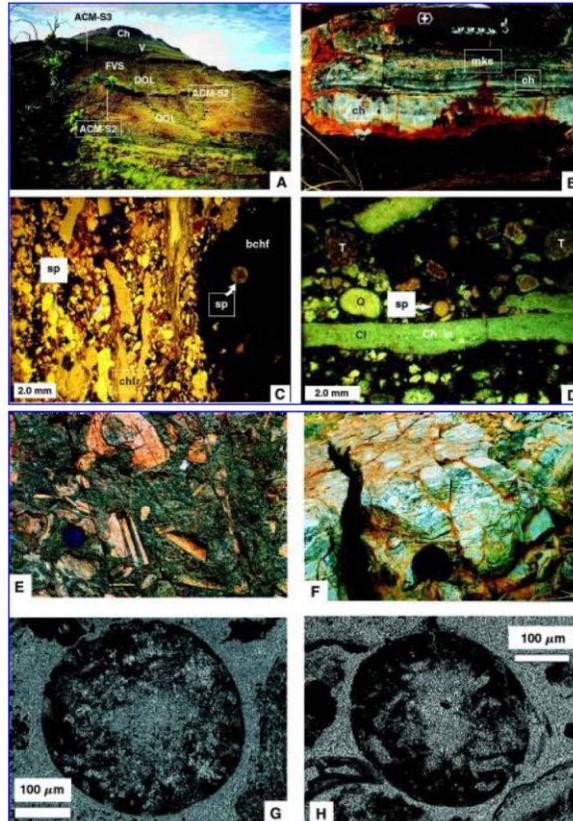


Fig. 1.2.2.5f. Debris flows, exotic boulders, and turbulence features associated with tsunamis triggered by an asteroid impact 3,470 mya in Pilbara Craton, Western Australia (*Astrobiology* 4, 19, 2004).

Even a very large impact does not directly affect life all over the Earth. However, its immediate consequences can easily be global. The emission of dust and particles causes environmental changes close to a nuclear winter, during which the surface of the Earth was totally covered by a cloud of dust for several years.

It seems very likely that extraterrestrial impacts were not the only triggers of drastic past changes in the environment and mass extinctions. Others triggers, not instantaneous but still very rapid at the geological time scale, were also probably important. Among those are episodes of unusually active volcanism (possibly responsible for the end-Permian extinction) and sudden changes in Ocean circulation, which could release huge amounts of H₂S and other toxic gases. Currently, causes of mass extinctions remain uncertain (Section 1.3.2.8).

Section 1.2.3. Direct evidence of past life and its evolution

A dead organism can easily decompose completely, and conversion of at least parts of its body into durable fossils, fossilization, occurs only under rather specific conditions. Fossils, studied by paleontology, come in a wide variety of ages, sizes, and types and represent hundreds of thousands of distinct kinds of organisms that lived from ~3,500 mya ago until the present time. Still, the fossil record is usually incomplete, and evolutionary reconstructions based on fossils require a number of inferences and should also use, if possible, information extracted from modern life.

1.2.3.1. Fossilization

Under most environments, fossilization requires unusual circumstances, and the vast majority of individuals leave to discernable traces of their existence. Fossilization normally depends on rapid burial of a dead organism, which restricts its contact with oxygen. Occasionally, fossilization follows freezing, desiccation, or trapping in a resin. Rarely, conditions favorable for fossilization, for example, lack of oxygen in a thick layer of mud, persist locally, leading to formation of Lagerstätten, sedimentary deposits with exceptional richness and preservation of fossils (Fig. 1.2.3.1a).





Fig. 1.2.3.1a. A "routine" path to fossilization (top left). "Lyuba", a female baby mammoth who died at the age of ~6 month and was preserved in permafrost for ~40,000 years in Yamalo-Nenetsk region of Russia (top right). Spider fossil in amber (bottom left). An assembly of crustaceans from the order Malacostraca found in Montceau-les-Mines Lagerstätte in France (bottom right) (*Journal of Paleontology* 83, 624, 2009).

Parts of dead organisms that escape decay undergo deep transformation before they reach states in which they can persist for billions of years. The original material is usually replaced by abiogenic minerals, although this process can still preserve even microstructural features. This process is analogous to slow diagenesis of sedimentary rocks that occurs at low temperatures and pressures. Empty spaces within a dead organism can also be filled by minerals that precipitate from the groundwater, a phenomenon known as permineralization (Fig. 1.2.3.1b).

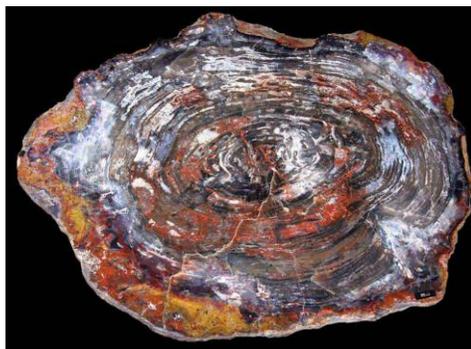
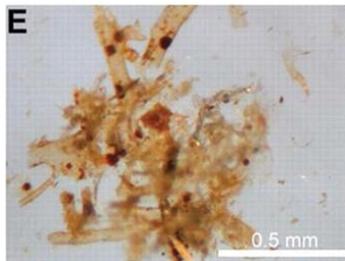


Fig. 1.2.3.1b. Cross-section of a petrified tree trunk, in which all the organic material have been replaced with silicated (left). A permineralized trilobite, *Asaphus kowalewskii* <http://en.wikipedia.org/wiki/Fossil> (right).

Naturally, different kinds of biological material have different potentials for fossilization. Hard, mineralized skeletons and shells of arthropods, vertebrates, mollusks and brachiopods can be fossilized easily. The same is true for microscopic shells of unicellular eukaryotes, such as radiolarians and foraminiferans, as well as for wood. In contrast, soft tissues are preserved only very rarely.

Fossils are mostly inorganic entities. Some very recent fossils, including those of Neanderthals and of woolly mammoth, harbor fragmented DNA which still can be used to determine their genome sequences. However, DNA disintegrates rapidly, and is unlikely to be found in fossils older than hundreds of thousands of years. In contrast, proteins may be more durable, and can be extracted from dinosaur bones. Sequences of dinosaur collagens confirmed close relatedness between birds and non-avian dinosaurs (Fig. 1.2.3.1c). Still, finding DNA or protein in a fossil is an exception, rather than the rule, although simpler organic compounds are found more often.



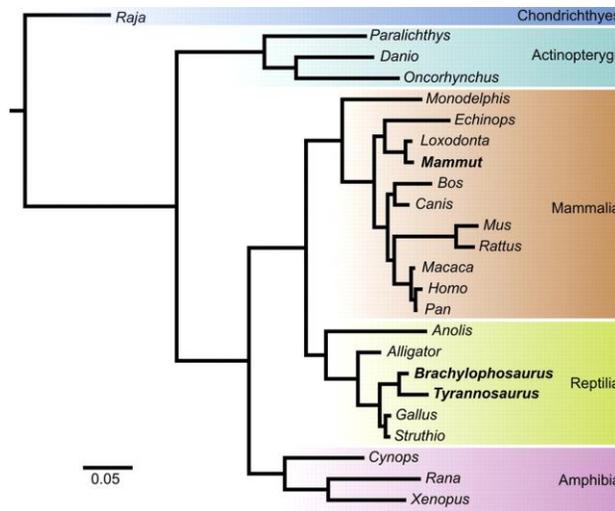


Fig. 1.2.3.1c. Micrograph of vessels and matrix in demineralized bones (left) of *Brachylophosaurus canadensis*, a 80-million-year-old dinosaur (reconstruction, right). A phylogeny, based on collagen sequences of three extinct (two dinosaurs and mastodon *Mammut*; boldfaced) and a number of extant species (bottom) (*Science* 324, 626, 2009).

Even a fully mineralized fossil can change further together with the sedimentary rock that harbors it. Thus, fossils are mostly found in rocks that did not undergo metamorphism. Fossils in metamorphic rocks of sedimentary origin may be disturbed beyond recognition. When a fossil reappears on the day surface, it also often decays fast. Of course, there are no fossils in igneous rocks.

1.2.3.2. Fossils

Despite the overall rarity of fossilization, the Earth still harbors a huge number of fossils of all kinds, reflecting a truly immense number of organisms that lived in the past. Currently, there are over 250,000 described fossil "species". Progress in paleontology, the science of fossils, to a large extent depends on simply finding more fossils, although new methods led to great improvement of the precision of age determination, as well as to discovery of fossils at cellular and even molecular levels. Fossils are found when erosion of the sedimentary rocks that hold them brings them back on the day surface, as well as in the course of excavations.

The most important kinds of fossils are body fossils that originate directly from parts of bodies of ancient organisms. Body fossils can be of various sizes and degrees of

completeness and preservation, and can consist of fossilized body parts, of prints left by them in sediments, or of permineralized body cavities (Fig. 1.2.3.2a). Trace fossils, preserved impressions made by organisms on substrates, are also common (Fig. 1.2.3.2b).

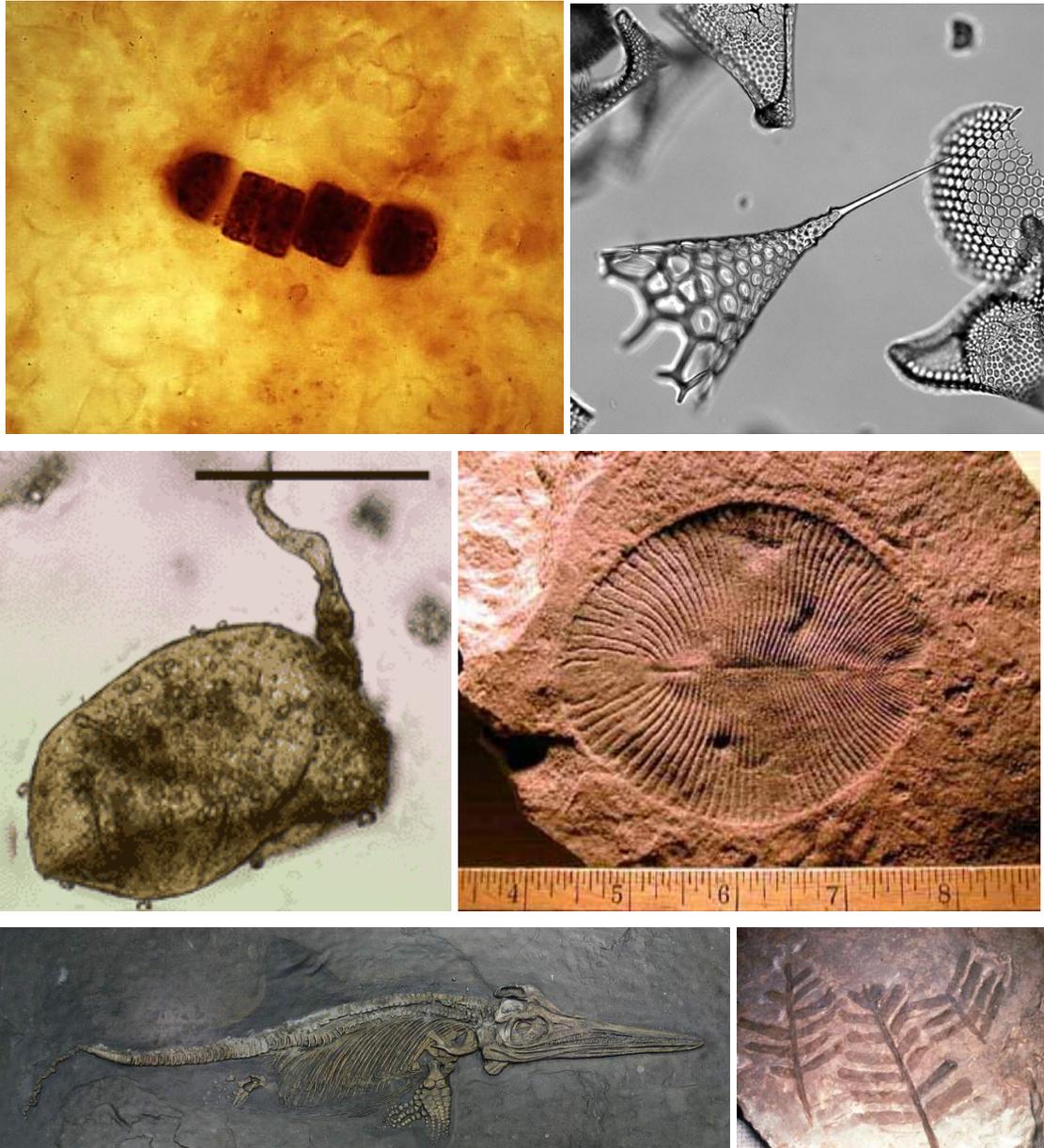




Fig. 1.2.3.2a. Body fossils. Bacterium from the Proterozoic eon (top left). 60 Ma old radiolarians, unicellular eukaryotes (top right; <http://www.teara.govt.nz/en/plankton/2/2>). Light microscope picture of a fossil fungal spore with attached hypha from the Ordovician, 460 mya (middle left; Dirk Redecker). A print fossil of unclear affinities *Dickinsonia*, found in the Ediacara Hills of Australia and near the White Sea coast in northern Russia (middle right). Ichthyosaur (bottom left). Fossil ferns from the Permian age Hermit Shale in Arizona's Grand Canyon (bottom right). 3.5 Ma old *Australopithecus afarensis*, very likely our direct ancestor (very bottom).



Fig. 1.2.3.2b. Trace fossils. Vertebrate traces in the Permian Coconino Sandstone, Arizona (left; <http://resweb.illu.edu/lbrand/research-bridger.html>). A fossilized dwelling burrow, probably of a marine crustacean from the Cretaceous Bad Heart Formation in Alberta (right; http://www.geo.ucalgary.ca/~macrae/t_origins/carbbones/burrow.html).

Modern methods of analysis can be applied to investigate fossils (Fig. 1.2.3.2c). However, interpreting fossils can still be difficult. For example, *Parapandorina* fossils from the Doushantuo region in China have features akin to those of animal embryos. The virtual cross-section reveals the component cells separated by walls. The cells' interiors are dominated by what seem to be membrane-bound vacuoles or yolk particles. However, *Parapandorina* could in fact be a bacterium similar to *Thiomargarita*, which exhibits some features superficially resembling animal embryos (Fig. 1.2.3.2d). Of course, many other fossils, especially of well-preserved macroscopic organisms similar to extant species, can be interpreted unambiguously.

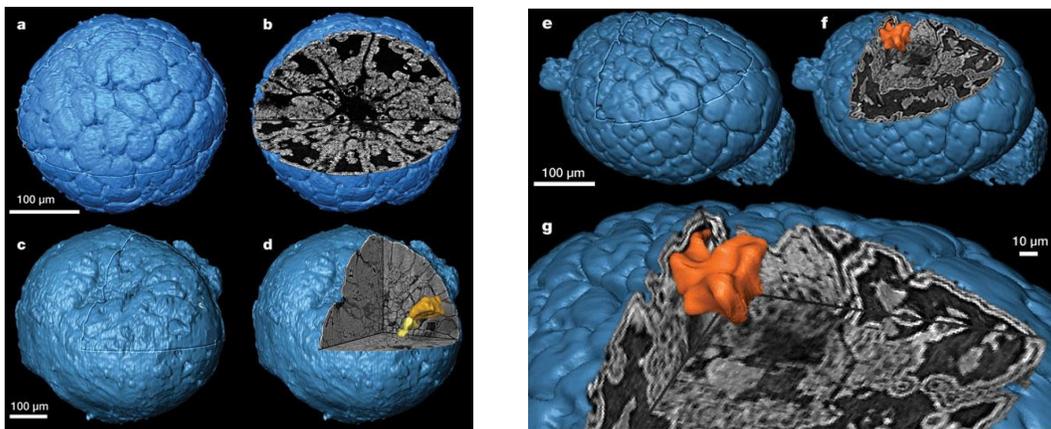


Fig. 1.2.3.2c. Tomographic reconstructions of cleavage-stage embryos from the Lower Cambrian Kuanchuanpu Formation, Southern China (*Nature* 442, 680, 2006).

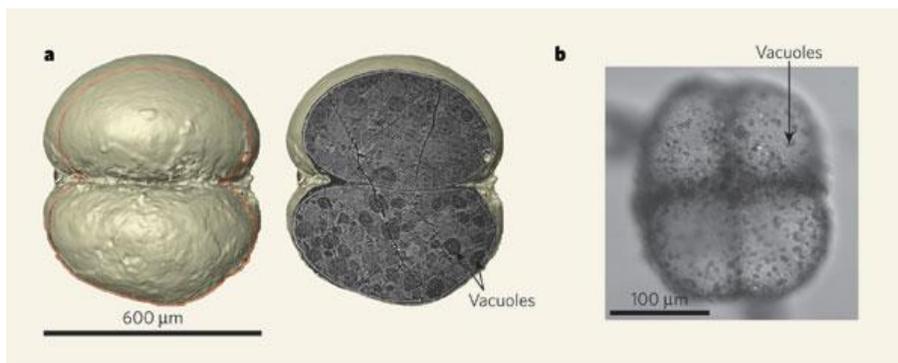


Fig. 1.2.3.2d. a, Image of a two-cell *Parapandorina* "embryo" obtained using X-ray microtomography. b, Extant bacterium *Thiomargarita*, divided into four cells, in which the vacuoles are clearly visible (*Nature* 445, 155, 2007).

Often, remnants of ancient organisms persist only as organic compounds, without any preservation of morphology. By far the most abundant and important fossils of this kind are fossil fuels, in particular, coal and oil. However, less abundant biogenic compounds, which constitute only a small fraction of sedimentary rocks, are also common. Often, such compounds are biomarkers, as they make it possible to identify, with some degree of precision, what kinds of organisms produced them. For example, long-chain fatty alcohols with more than 23 atoms are a biomarker of planktonic bacteria.

1.2.3.3. Fossil record

Fossil record of past life is, in two ways, very far from being complete. First, only a small fraction of kinds of organisms that existed in the past are represented by known fossils, and in many groups fossilization is very rare. Second, fossils preserve information on only a small fraction of the phenotypes of ancient organisms and usually reveal nothing about their genotypes.

In terms of the numbers of fossilized individuals, the most complete fossil records are those of plankton marine microorganisms with hard shells, such as coccolithophores, whose fossils produce chalk, radiolarians, and foraminifera. It is not unusual to have, in the Ocean bottom deposits, essentially continuous records of these organisms that cover many millions of years and are represented, for each moment, by an essentially limitless number of specimens. Such records make it possible to observe gradual changes of traits during anagenesis (Fig. 1.2.3.3a) and, rarely, events of cladogenesis (Fig. 1.2.3.3b).

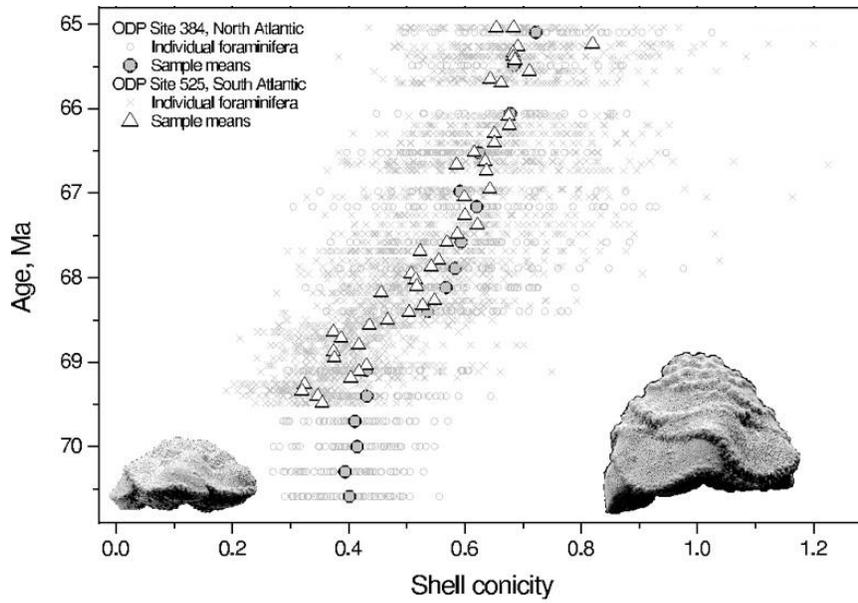


Fig. 1.2.3.3a. Gradual increase through time in shell conicity of the planktonic foraminifer *Contusotruncana* at two distant locations in the Late Cretaceous Atlantic Ocean (*PNAS* 105, 12382, 2008).

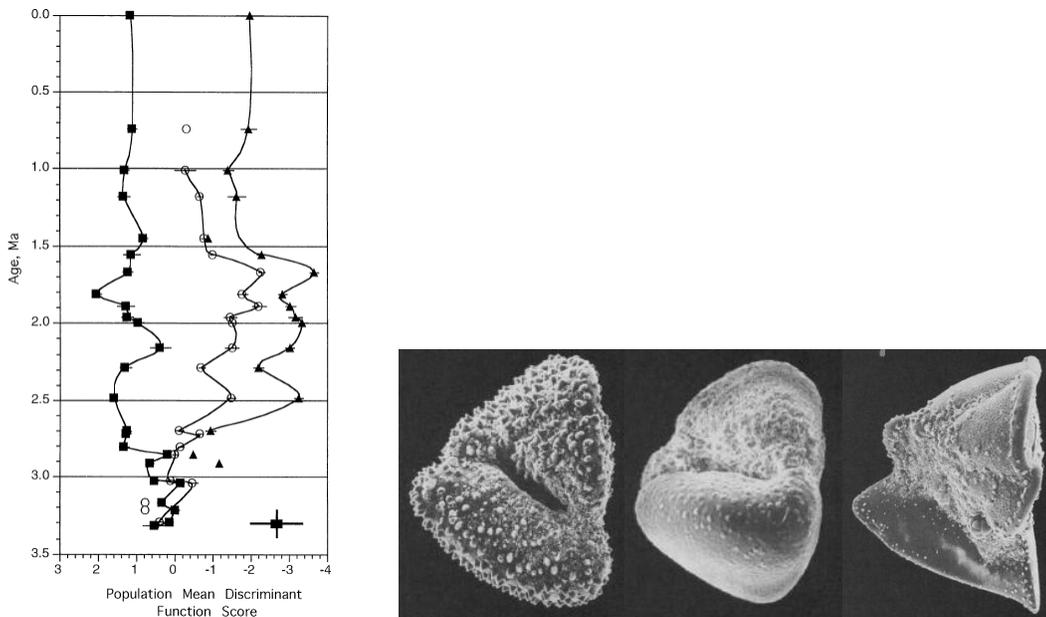


Fig. 1.2.3.3b. Two nearly simultaneous cladogeneses that produced three foraminifer species, *Globorotalia crassiformis* (squares, left image), *G. tosaensis* (circles, center image), and *G. truncatulinoidea* (triangles, right image), in the course of the last 3.5 Ma.

Points are population means scored on a discriminant function separating all measured specimens of different species (*Paleobiology* 21, 28, 1995).

However, incomplete preservation of information may lead to serious problems in interpretation of even such continuous fossil records. In particular, traits preserved in fossils may be insufficient to distinguish closely related species, and what seems to be a single ancestral-descendant lineage (Fig. 1.2.3.3a) may, in fact, be a bunch of genetically independent lineages that evolve simultaneously, perhaps affecting each other due to competition. The courses and outcomes of evolution may be very different in these two situations.

At the other extreme are fossil records of large organisms, where only a very limited number of specimens is usually known. For example, so far there are only 7 known fossils, neither of which with a complete preservation of even the skeleton, pertinent to the evolution of terrestrial tetrapods from sarcopterygian fishes, which lasted from 385 until 360 mya, despite enormous efforts to shed light on this key stage of the evolution of vertebrates (Fig. 1.2.3.3c; Chapter 1.3). For many crucial evolutionary transitions, such as the origin of eukaryotes, vertebrates, or flowering plants, there is no known fossil record. In some cases, this may change in the future, because new discoveries are constantly being made, but some evolutionary transitions, unfortunately, may really leave no fossil record.

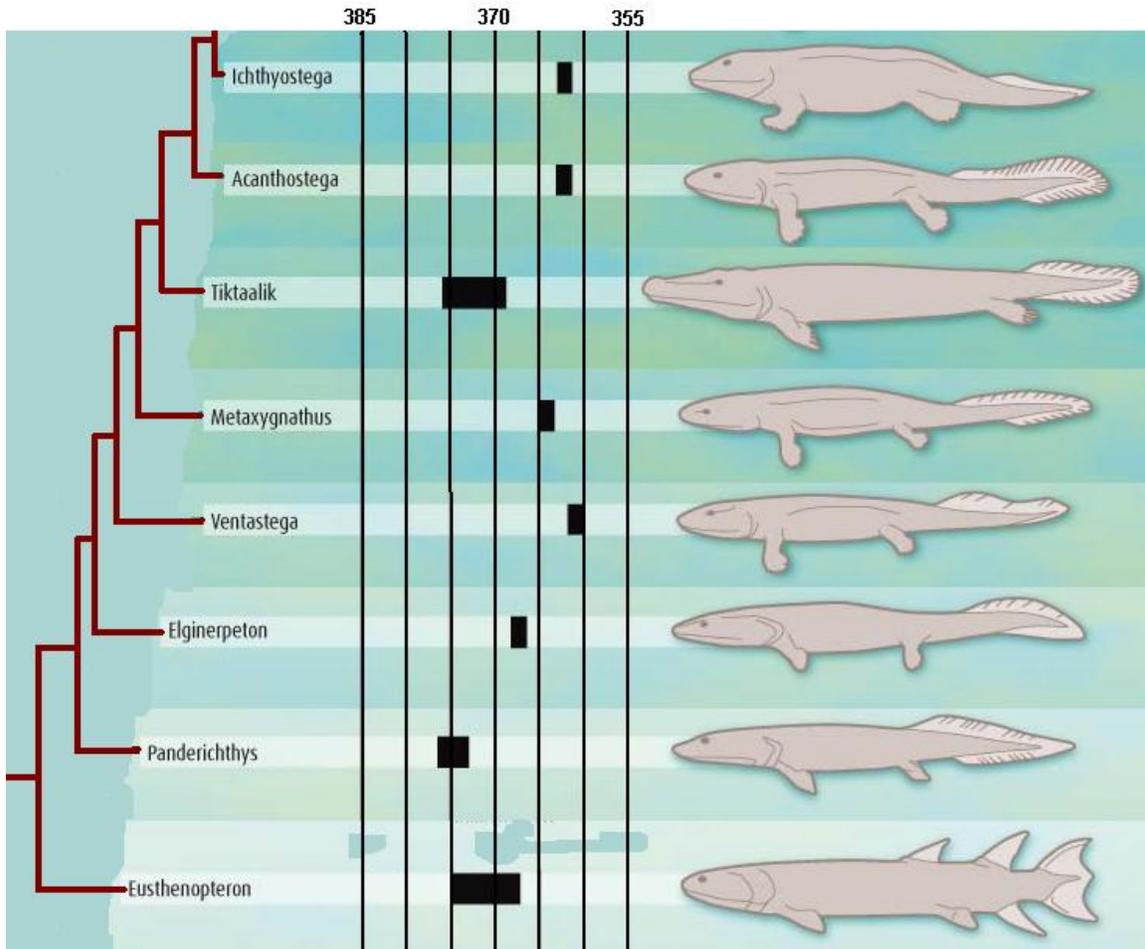


Fig. 1.2.3.3c. Known fossils that span the gap between sarcopterygians and amphibians (<http://www.newscientist.com/data/images/archive/2645/26451701.jpg>).

When the fossil record is complete enough to reveal the course of evolution of a quantitative trait within a lineage, the rate of this evolution needs to be characterized. This can be done in at least two ways. First, we can assume that the rate of evolution r of the mean value of a trait x is proportional to its current value: $dx/dt = rx$ (of course r can be both positive and negative, and this equation is identical to 1.2.2.1a). Then, the rate of evolution is characterized by r , and $r = Ma^{-1}$ is called a darwin. Because $x(t) = x_0e^{rt}$, (Section 1.2.2.1) the mean value of a trait evolving at a rate of 1 darwin will change by a factor of $e = 2.718\dots$ each Ma. Second, the rate of evolution of the mean value of a quantitative trait can be measured in the units of its standard deviation within the evolving population (Chapter 2.1). A haldane is a rate of evolution such that the mean value of a trait changes by one standard deviation each generation. Rate of evolution of

more complex phenotypes can also be characterized, for example, by how often qualitatively new trait states appear in the fossil record.

1.2.3.4. Evolutionary inferences from fossils

Fossils are direct evidence of past life and, to the extent they reveal its evolution, direct evidence for past evolution. Indirect and direct data on past life complement each other remarkably well. It is possible to at least partially infer ancestral genomes by comparing modern genomes, but doing the same with phenotype-level traits is, at least now, almost impossible. In contrast, DNA is poorly preserved in fossil record, but some organism-level traits are often preserved very well. Also, fossils are the only source of information on extinct forms of life which left no survivors: without paleontology, we would never know of *T. rex*. Precise timing of many crucial transitions in past evolution is also primarily revealed by fossils.

Because records of anagenesis are very incomplete, and records of cladogenesis are exceptionally rare (Fig. 1.2.3.3a,b), the first serious problem in evolutionary interpretation of fossils is establishing ancestor-descendant relationships between fossils of different ages, or between past and extant organisms. Here, caution is needed. Even when the fossil record is continuous, there may be problems due to cryptic species, and when we have to deal with sparse fossil records, revealing ancestor-descendant relationships becomes very problematic. It is unlikely, for example, that *Tiktaalik* was a direct ancestor to *Ichthyostega* (Fig. 1.2.3.3c), although this possibility is not ruled out by the available data. More likely, *Tiktaalik* and *Ichthyostega* represent separate lineages, produced by cladogenesis shortly before the time when *Tiktaalik* lived.

However, in some important cases direct ancestor-descendant relationships between known organisms that lived at different times are plausible. For example, it is very likely that *Australopithecus afarensis* (Fig. 1.2.3.2a) represents a segment of the lineage that led to modern *Homo sapiens*, and thus, is our direct ancestor. Indeed, recent evolutionary history of hominids is elucidated by a considerable number of fossils, and it is unlikely that ~3.5 mya there lived another lineage of hominids, closely related to *A. afarensis*, which eventually gave rise to *H. sapiens*. Still, this cannot be ruled out

currently, and there were evolutionary dead-ends in the history of hominids, such as extinct *Paranthropus*, which could be mistaken for a human ancestor (Chapter 1.4).

Even a fossil which is not a direct ancestor of a particular clade, extinct or extant, may still shed light on the course of evolution of this clade. For example, *A. afarensis* possessed bipedal locomotion but its brain size was still quite small (Chapter 1.4). On the basis of this fact we can conclude that in the course of human evolution bipedal locomotion evolved before large brain, because even if *A. afarensis* did not belong to the lineage of *H. sapiens*, it must have been very close to it. As a rule, we observe an excellent correspondence between the order in which different traits appear in the fossil record and their distribution relative to each other among extant organisms: hierarchically subordinate trait states appear later (Fig. 1.2.3.4a). This correspondence may be viewed as yet another kind of evidence for past evolution, and it demonstrates utility of fossils for reconstructing its course.

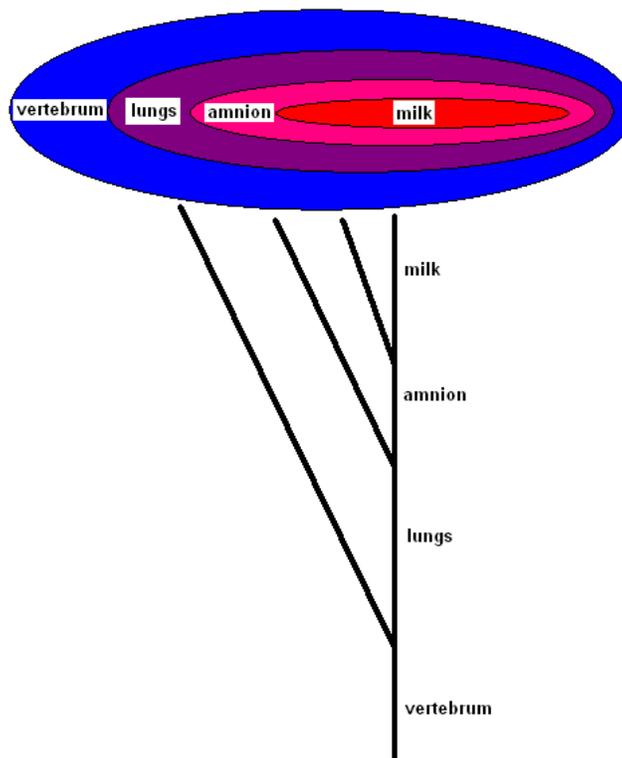


Fig. 1.2.3.4a. A correspondence between the timing of the first appearance of a trait state in the fossil record and its position within the hierarchy of traits of extant organisms.

Three concepts are very useful for describing relationships between fossil and extant organisms. Let us consider all living species that belong to a particular clade, for example, all extant birds. Then, all organisms, extant and extinct, which originated from the most recent common ancestor of extant birds are said to belong to the crown group of birds. They include both direct ancestors of some modern bird species and extinct forms that left no descendants. Extinct organisms that diverged before the last common ancestor of living birds but still are more tightly related to birds than to any other living organisms are said to belong to the stem group of birds. Of course, members of the stem group of birds are also members of the crown group of vertebrates (Fig. 1.2.3.4b). Together, the crown and the stem groups of a clade of living species comprise its total group.

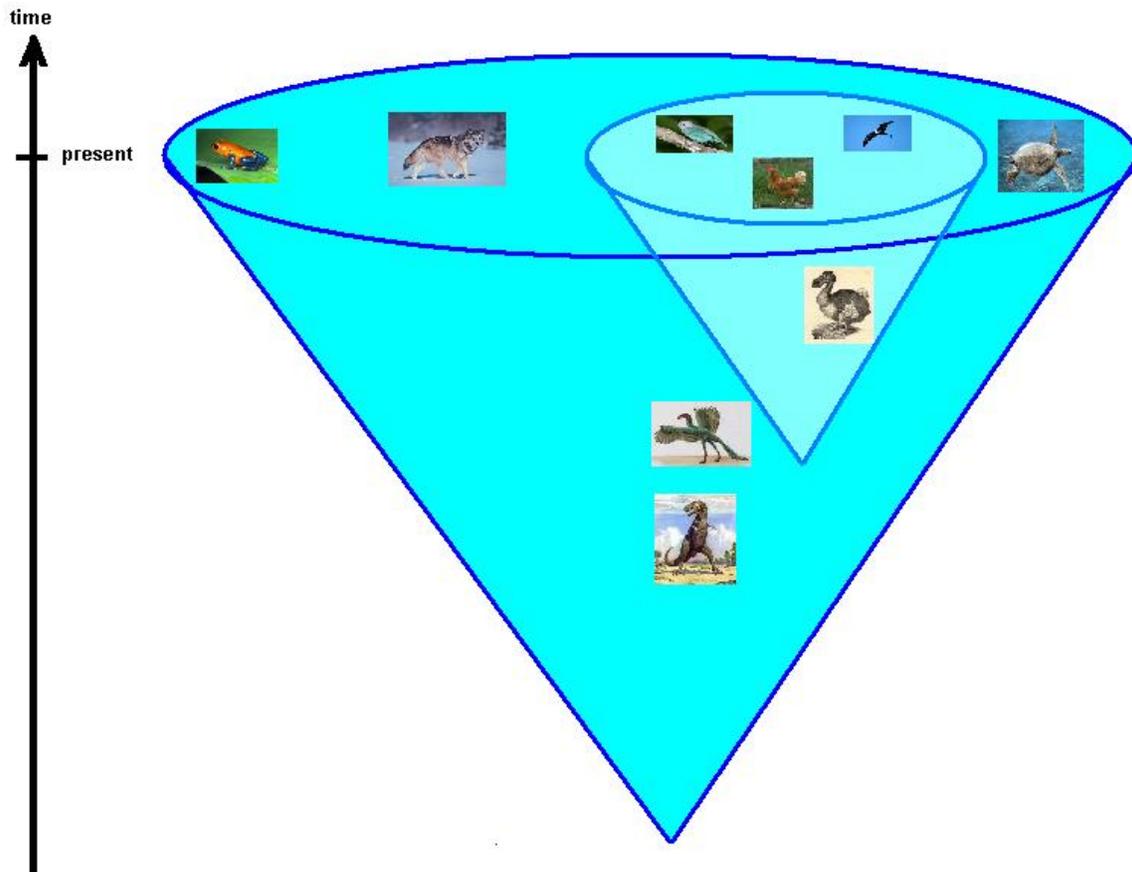


Fig. 1.2.3.4b. Extinct *Dodo* belongs to the crown group of living birds – it originated from their most recent common ancestor. *Archaeopteryx* and *Tyrannosaurus* belong to the stem group of living birds – they did not originate from their most recent common ancestor but

share a more recent common ancestor with living birds than with any other living organism – and to the crown group of all tetrapods (Fig. 1.2.3.1c).

Section 1.2.4. History of abiotic conditions on Earth

During much of its history, life on Earth experienced conditions very different from what we see today. Periods of globally warm climates alternated with Ice Ages, including the current one. Positions of continents constantly changed, and there were at least three long periods where landmasses mostly constituted just a single supercontinent, the last of which, Pangea, broke up ~200 mya. Starting from ~2,500 mya, enormous amounts of oxygen, produced by cyanobacteria, were absorbed by Fe^{2+} that was dissolved in the oceans, and oxygenation of deep ocean waters occurred only ~2,000 Ma later. Oxygen began to accumulate in the atmosphere 2.400 mya but did not approach its current concentration until 542 mya. Taken together, the data obtained by the Earth sciences, including paleontology, established a geological time scale which provides the necessary background for studying past evolution.

1.2.4.1. Climate

In the past, global climate underwent many very deep changes. Among the causes of these changes are Earth orbital variations, continental drift, changes in the Sun luminosity, changes in volcanism (caused by plate tectonics), changes in the atmosphere composition, and extraterrestrial impacts. As the result, periods of warm temperatures alternated with Ice Ages, where a substantial fraction of Earth was covered by ice (Fig. 1.2.4.1a). Neoproterozoic glaciations can be recognized on the basis of C- and Sr-isotope profiles in underlying and overlying strata and distinctive cap carbonates and were probably triggered by the breakup of Rodinia, which changed circulation of oceanic waters. The Sturtian and the Marinoan were perhaps the greatest glaciations Earth ever experienced, with reliable paleomagnetic evidence of glaciers at sea level within 10° of the equator (*Precambrian Res.* 111, 1, 2001).

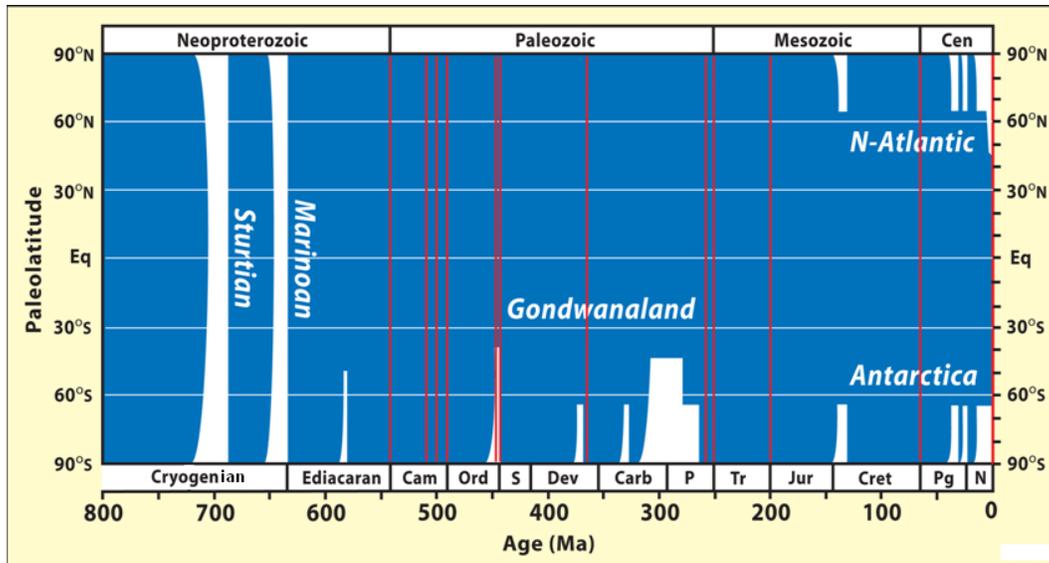


Fig. 1.2.4.1a. Paleogeographic extent of continental ice sheets and permanent sea ice over the last 800My (red lines indicate major mass extinctions).

Well-established Ice Ages include:

- 1) Huronian, 2,400-2,100 mya;
- 2) Cryogenian, 850-630 mya. This Ice Age included two major glaciations, Sturtian (725-710 mya) and the Marinoan (660-635 mya);
- 3) Andean-Saharan, 460-430 mya;
- 4) Karoo, 330-260 mya;
- 5) Pliocene-Quaternary, which started 2.58 mya and continues until the present, with glaciations alternating with periods of higher temperatures.

Outside these Ice Ages the Earth was mostly ice-free at all latitudes. However, there were also isolated Glacial Periods, including the Gaskiers (580Mya), which may have triggered the evolution of Ediacaran biota (Section 1.3.1.4). Complementarily, there were also thermal maxima between Ice Ages (Fig. 1.2.4.1b), which likely led to episodes of widespread anoxia in the ocean (Section 1.2.4.3) and at least to some of the mass extinctions (Section 1.3.2.8).

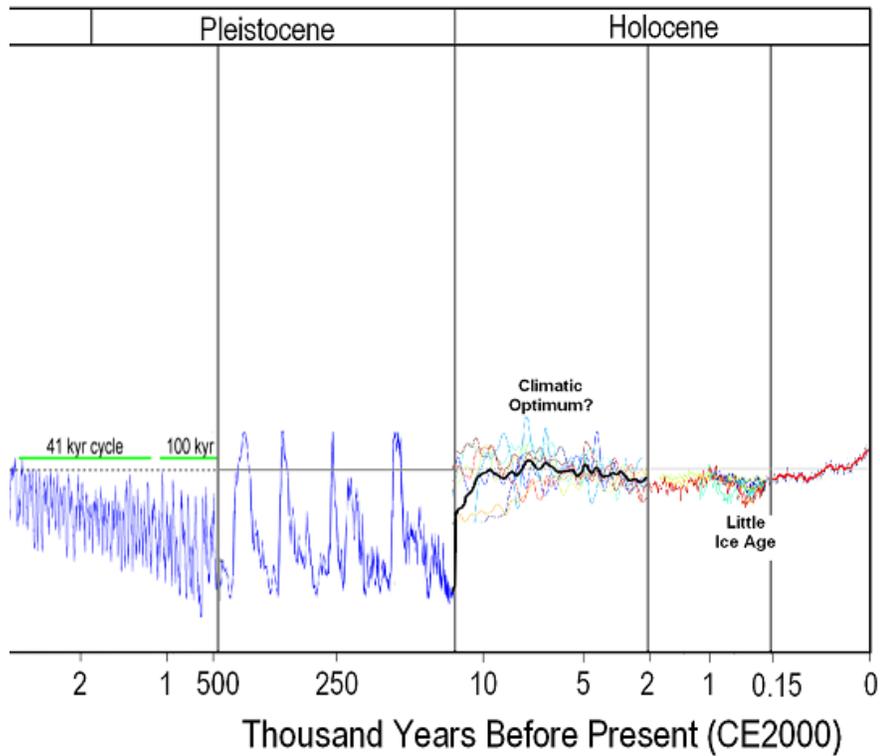
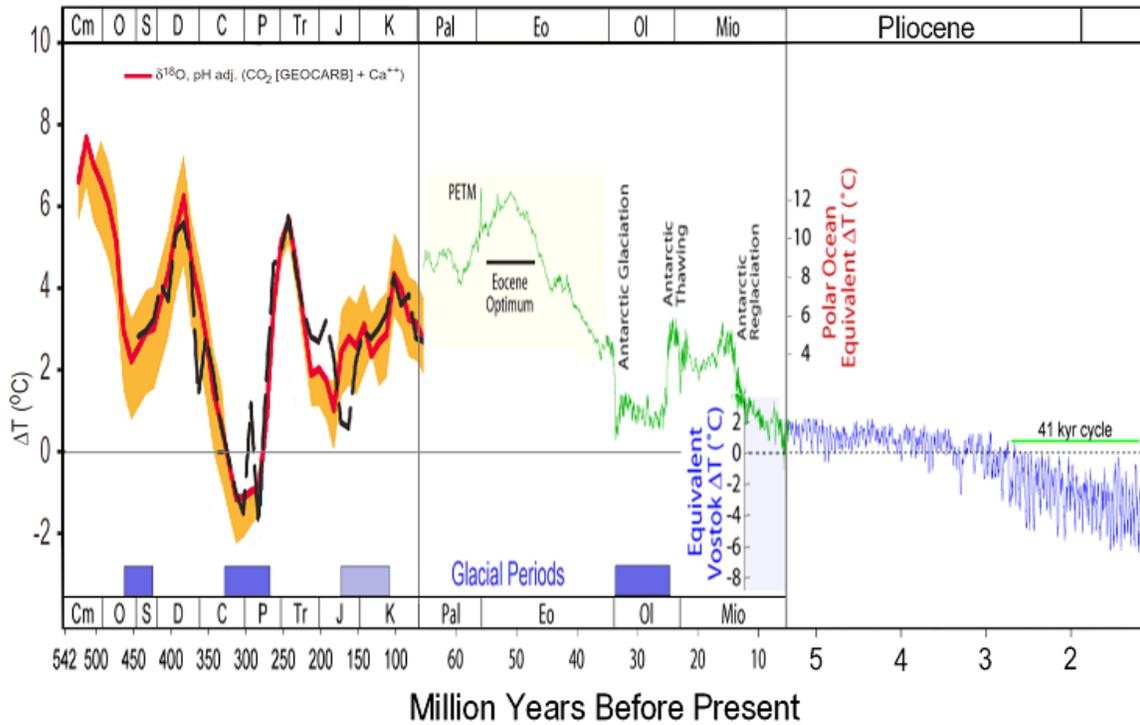


Fig. 1.2.4.1b. Average global temperature, relative to the current one, estimated through $\delta^{18}\text{O}$, during the Phanerozoic eon (<http://en.wikipedia.org/wiki/Paleoclimatology>).

1.2.4.2. Geography

A variety of evidence made it possible to reconstruct past changes of the geography of the Earth at the global scale. Although many details remain obscure, several key steps in these changes are known with a good measure of confidence:

1) ~1100 mya, most of Earth's landmass formed one supercontinent, named Rodinia. Rodinia broke up ~750Mya, probably into three continents.

2) ~600 mya, a single supercontinent Pannotia was formed again, to break up 540 mya, probably into four continents.

3) ~300 mya, a single supercontinent Pangea (Fig. 1.2.2.3b) appeared again.

4) The break-up of Pangaea began ~200 mya, first into two supercontinents (Gondwana to the south and Laurasia to the north), thereafter into the continents that exist today (Fig. 1.2.4.2a).

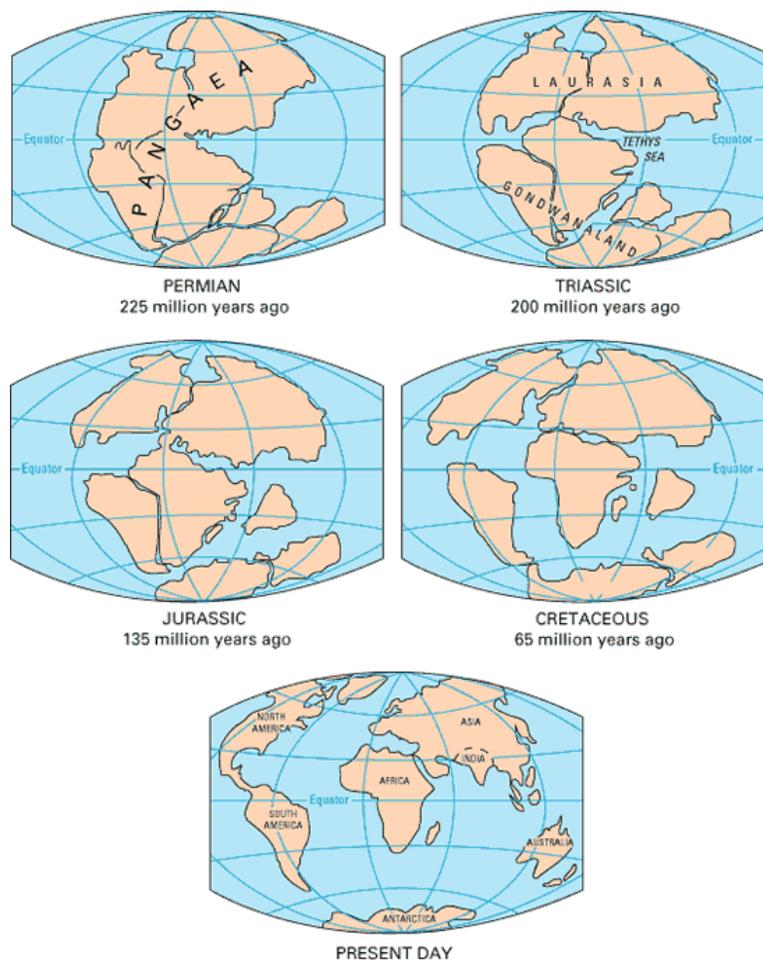


Fig. 1.2.4.2a. Changes in the locations of continents during the last ~200 Ma.

1.2.4.3. Oceans

Because early life mostly evolved in the sea, properties of ocean waters are of particular importance. During the history of the Earth, there were times when ocean waters contained high concentrations of dissolved Fe^{2+} (ferruginous conditions), H_2S (sulfidic, or euxinic conditions), and O_2 (oxic conditions). To complicate the matters further, the ocean was often stratified, with different layers of water having very different properties (today, the ocean is oxic from top to bottom).

Although oxygenic photosynthesis, the only source of O_2 , apparently evolved before ~2,700 Ma ago (Section 1.3.1.2), oxygen did not begin to accumulate in the ocean until much later. The top layer of the ocean water became oxygenated to some extent after ~2,400 mya, simultaneously with initial oxygenation of the atmosphere, and there was apparently no oxygen in deep water until ~575 mya. Initially, the oxygen released by cyanobacteria reacted with Fe^{2+} dissolved in the oceans, to form insoluble iron oxides Fe_2O_3 and Fe_3O_4 . These oxides were deposited on the ocean floor as banded iron formations (BIFs; Fig. 1.2.4.3a). Accumulation of BIFs was very common between ~2,500Mya and ~1,800Mya, after which it almost stopped for ~1000 Ma (Fig. 1.2.4.3b). Most of old BIFs show microbanding, interpreted as annual layers of sedimentation. The chemistry of BIFs, from 3,000My through 1,800My in age, is very similar. They are rich in Fe (20-40%) and SiO_2 (43-56%) and are probably clean sediments devoid of detrital input.



Fig. 1.2.4.3a. The Banded Iron Formation from Soudan Mine, MN. Deposited at the sea bottom 2,700 mya, it was folded by the movements of plates.

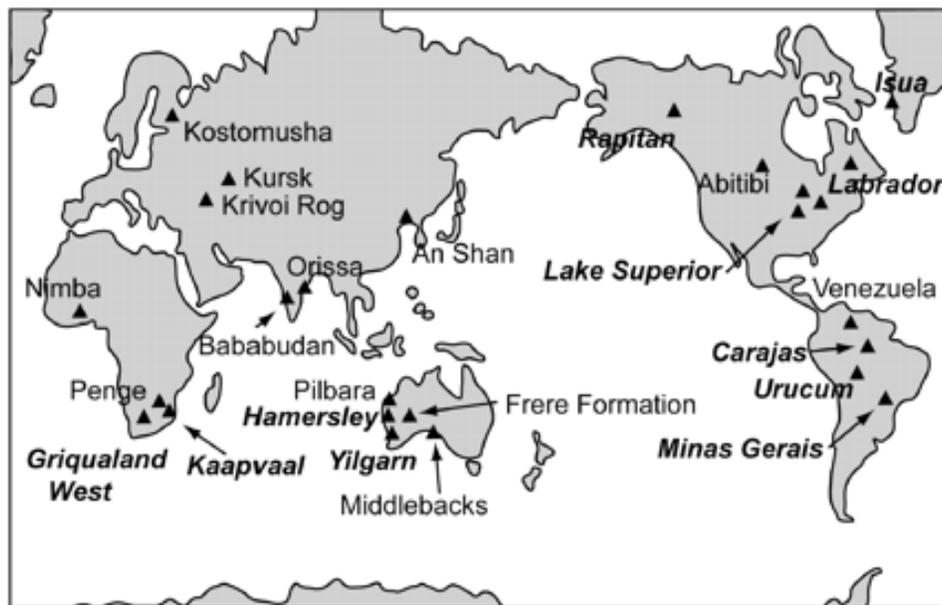


Fig. 1.2.4.3b. Global occurrence of BIFs. The map emphasizes the wide distribution of BIFs and shows only a selection of well-known occurrences.

There is an intriguing possibility that some large fragments of heavy BIFs submerged through the Earth mantle and are currently floating at the surface of its core, being observed as the so-called ultralow-velocity zones, regions of the Earth's core-mantle boundary $\sim 1\text{--}10\text{km}$ thick exhibiting seismic velocities that are reduced by

10–30% (Fig. 1.2.4.3c). If this hypothesis, which explains all the observed features of ultralow-velocity zones, is correct, ancient submerged BIFs provide the only known example of the influence of life reaching so deep inside the Earth.

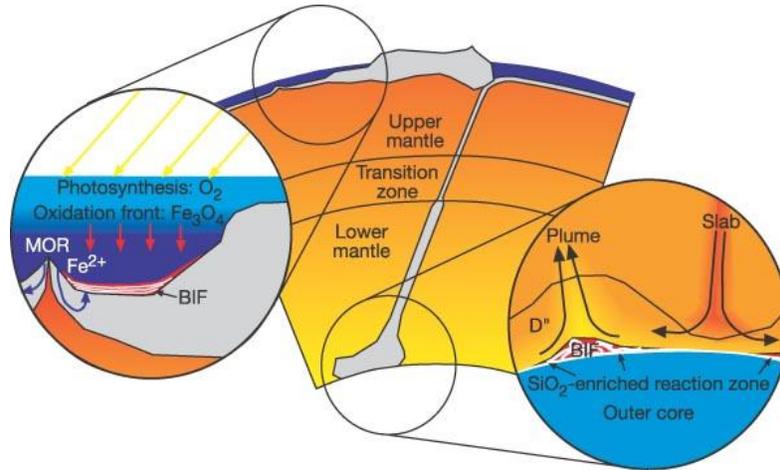


Fig. 1.2.4.3c. Left inset, early life in the photic zone of the oceans releases oxygen into the atmosphere by photosynthesis. This reacts with Fe^{2+} in deep ocean waters to produce insoluble iron oxides, which are precipitated as BIF on the ocean floors. BIF is recycled into the mantle at subduction zones. MOR, mid-ocean ridge. Right inset, BIF ponds at the core–mantle boundary, where it is neutrally buoyant. Convection currents concentrate BIF below regions of mantle upwelling where the ultralow-velocity zone is now observed (*Nature* 434, 371, 2005).

The first period of deposition of BIFs probably ended due to transition of deep ocean water to sulfidic state. Apparently, the increasing atmospheric oxygen levels enhanced sulphide weathering on land and the flux of sulphate to the oceans. As a result, between 1,800 and 800 mya deep oceanic water contained a lot of H_2S and almost no O_2 , as in the Black Sea today, making further formation of BIFs impossible. However, deposition of BIFs resumed between 800 and 600 mya. These younger BIFs had mineralogical and chemical make-ups different from those 2,500-1,800 mya BIFs, and were probably the result of mostly anoxic conditions due to stagnation in the oceans beneath a near-global ice cover.

Fig. 1.2.4.3d summarizes the current state of knowledge of past compositions of oxygen minimum zone (OMZ, a layer of oceanic water below surface waters, that became

oxygenated ~2,400 mya), and of factors that affected them. Persistence of sulfidic deep water during ~1000 mya between the 1st and the 2nd periods of accumulation of BIFs was likely due to prevalence of anoxygenic photosynthesis at that time, and eventual oxygenation of these waters was probably triggered by changed patterns of circulation, eventually leading to predominance of oxygenic photosynthesis by eukaryotic algae and sea weeds and to diversification of heterotrophic eukaryotes (Section 1.3.1). However, these conclusions are only tentative. Still, we know that oxygenation of the deep oceans lagged that of the atmosphere by ~1,000 Ma (*Nature Geoscience* 1, 597, 2008).

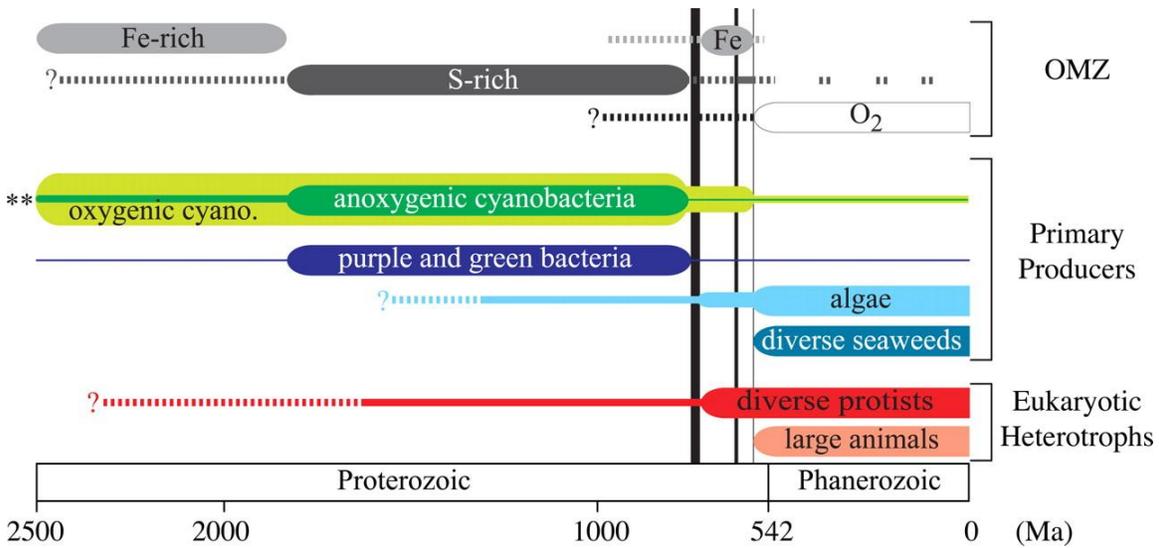


Fig. 1.2.4.3d. A timeline showing OMZ chemistry, the relative contributions from different primary producers, and the evolution of eukaryotic heterotrophs. Band thicknesses approximate the importance of each feature through time. Dashed lines represent postulated or uncertain histories. The two thicker vertical lines represent the major Cryogenian glaciations, Sturtian and Marinoan, and the thinner line to the right represents the Ediacaran Gaskiers glaciation. Neoproterozoic climatic and biogeochemical events remains a subject of ongoing research (*Science* 321, 949, 2008; *PNAS* 106, 16925, 2009).

However, a likely oxygenation of OMZ after Gaskiers glaciation probably was not followed by complete oxygenation of the ocean. Instead, there are evidence that sulfidic conditions persisted in deep water, and that its upwelling at the end of the Ediacaran

period led to temporary anoxia in even the top layer of water (Section 1.3.1.4). In the Phanerozoic eon, the ocean was mostly oxygenated most of the time, including deep water, but episodes of widespread and even global sulfidic conditions (OAEs, oceanic anoxic events) in deep waters also occurred. These episodes may be triggered by increases in the global temperature, because of reduced solubility of oxygen in warmer water. In particular, there were apparently global sulfidic conditions, at least in deep ocean waters, during periods of high temperature in the Devonian, Permian, and Cretaceous periods (Fig. 1.2.4.1b; *Annual Review of Earth and Planetary Sciences* 36, 251, 2008).

1.2.4.4. Atmosphere

Changes in the concentration of oxygen in the atmosphere not only deeply affected evolving life but were also caused to a large extent by the activity of living beings. Presumably, the atmosphere of the primordial Earth mostly consisted of H_2 and He. These gases were rapidly lost into space, because the Earth's gravity is not strong enough to hold them, and replaced by volcanic gases. Still, atmosphere remained oxygen-free until $\sim 2,400$ mya (Fig. 1.2.4.4a), although oxygenic photosynthesis, the only source of atmospheric O_2 , evolved 300-500 Ma earlier (Section 1.3.1.2). Between 2,400 and 1,850 mya, concentration (partial pressure) of oxygen in the atmosphere rose from 0% to $\sim 5\%$ (this process is often called Great Oxidation Event), after which it probably did not change much for another 1,000 Ma, and exceeded 10% only by the end of the Neoproterozoic. In the course of the Phanerozoic, concentration of O_2 varied between $\sim 35\%$ in the late Paleozoic and $\sim 10\%$ in the early Mesozoic.

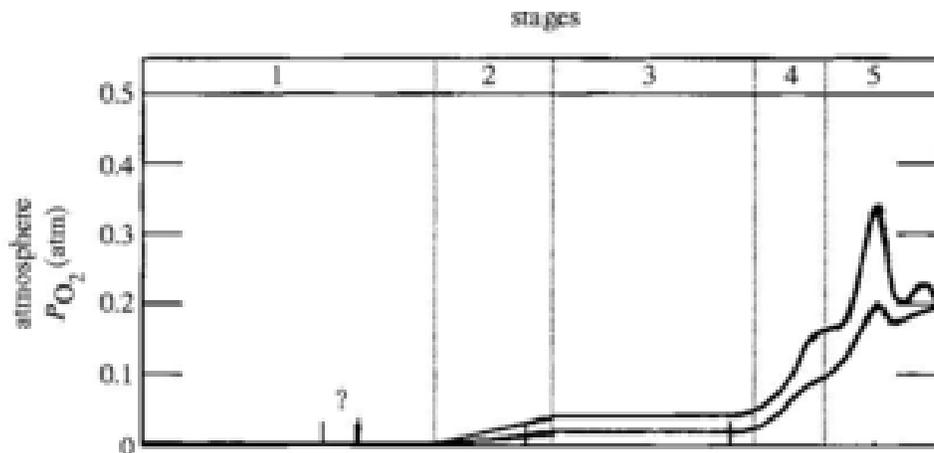


Fig. 1.2.4.4a. O₂ build-up in Earth's atmosphere (upper and lower curves represent the range of estimates): 1) 3,850–2,400Mya: no O₂, 2) 2,400–1,800Mya: O₂ produced but reached only a low concentration, being mostly absorbed in oceans and seabed rock, 3) 1,800–0,800Mya: O₂ produced but mostly absorbed by land surfaces and formation of ozone layer, 4) 800–540Mya: O₂ sinks filled and the gas accumulates, 5) 540 mya – present: near-current concentrations of O₂ (*Processes on the Early Earth* 405, 33, 2006).

Concentration of CO₂ also changed substantially in the past, affecting global climate due to greenhouse effect. The end of Cryogenian Ice Age may be caused a a rapid increase of the CO₂ concentration, perhaps due to global release of methane from low-latitude permafrost (*Nature* 453, 642, 2008). More reliable inferences on CO₂ concentration are currently available for the Cenozoic era (Fig. 1.2.4.4b).

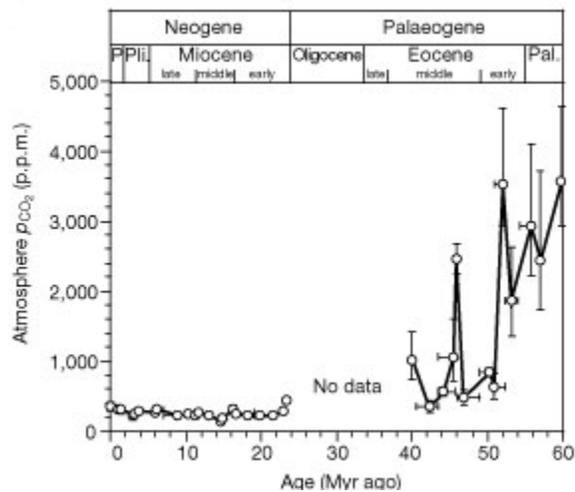
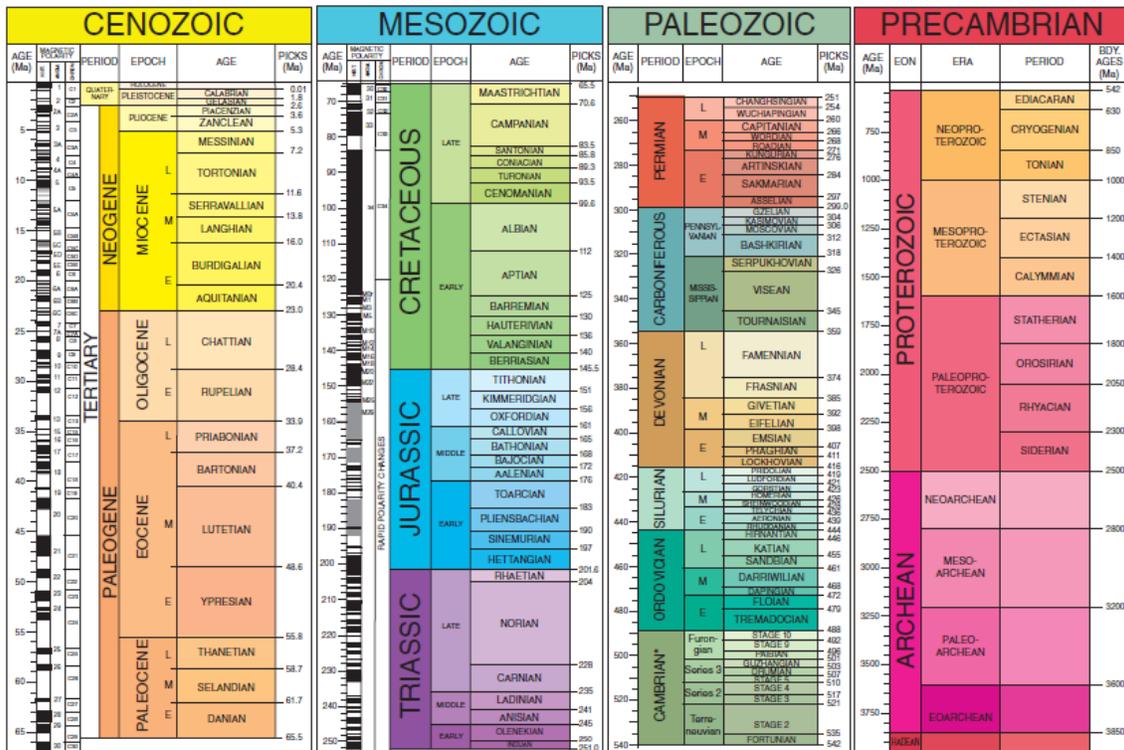


Fig. 1.2.4.4b. Atmospheric CO₂ concentrations were above 2,000 p.p.m. from 60 to 52 mya and declined between 55 and 40 mya, perhaps due to reduced CO₂ outgassing from ocean ridges and volcanoes and increased carbon burial. Since 24 mya, atmospheric CO₂ concentrations appear to have remained below 500 p.p.m. and were more stable than before. These conclusions are based on estimates of the pH of surface-layer sea water from the δ¹¹B in fossil planktonic foraminifer shells (*Nature* 406, 695, 2000).

1.2.4.5. Geological time scale

Data on past changes of non-living and living matter on Earth can be summarized in what is known as Geological time scale. Because long intervals of gradual changes were alternating with episodes of rapid, drastic changes, the history of the Earth can be naturally subdivided into separate entities: eons, eras, periods, etc. Boundaries of periods in Phanerozoic eon are defined mostly by abrupt changes in the fossil record. Geological time scale is traditionally shown with time running up, reflecting the arrangement of strata of sediments, but a horizontal arrangement may also be convenient (Fig. 1.2.4.5a).

2009 GEOLOGIC TIME SCALE



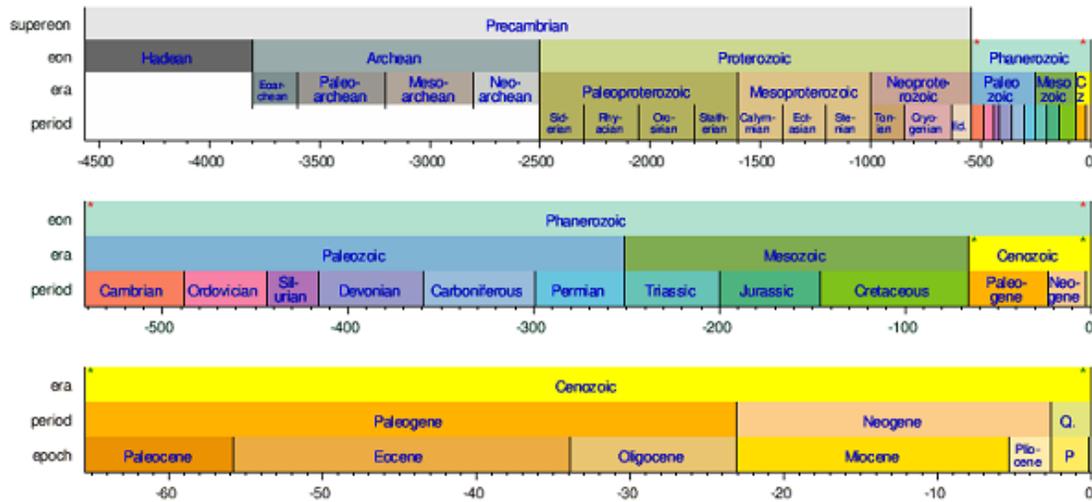


Fig. 1.2.4.5a. Two drawings of the Geological time scale (for more legible pictures go to <http://www.geosociety.org/science/timescale> and http://en.wikipedia.org/wiki/Geologic_time_scale).

Geological time scale provides the universal framework to organize data about past, concerned both with changes of non-living matter and with the evolution of life. Below, I summarize these data, presented in this and in the following Chapter. Data on processes and phenomena that took place over a long period of time are presented according to the times when they begun.

KEY FACTS ABOUT HISTORY OF THE EARTH AND LIFE

ARCHEAN eon: 3,850 - 2,500

Before 3,500: Origin of life

3,466-3,430: Earliest fossils of prokaryotes, including stromatolites

2,900: Likely origin of cyanobacteria and oxygenic photosynthesis

2,600: The first likely traces of terrestrial prokaryotes

PROTEROZOIC eon: 2,500 - 542

PALEOPROTEROZOIC era 2,500 - 1,600

2,500-1,840: Massive accumulation of Banded Iron Formations (BIFs)

2,400-2,320 Substantial accumulation of previously absent O₂ in the atmosphere

2,400-2,100: Huronian Ice Age

- 2,023: Vredefort, the oldest and largest (diameter ~300km) impact crater known
- 1,800: The first plausible fossils of unicellular eukaryotes
- 1,800-800: Probably, high concentrations of H₂S in deep oceanic waters

MESOPROTEROZOIC era 1,600 - 1,000

- 1,300: Abundant fossils of eukaryotes, including some multicellular forms
- 1,100-750: Supercontinent Rodinia

NEOPROTEROZOIC era 1,000 - 542

Tonian period: 1000-850

- 850: Diversification of multicellular eukaryotes

Cryogenian period: 850-635

- 800-600: Second period of accumulation of BIFs
- 800-640: Initial diversification of animals
- 725-710: Sturtian glaciation
- 660-635: Marinoan glaciation

Ediacaran period: 635-542

- 600-540: Supercontinent Pannotia
- 580: Gaskiers glaciation
- 575-542: Ediacaran biota, the first record of diverse macroscopic animals

PHANEROZOIC eon: 542 - present

PALEOZOIC era: 542-251

Cambrian period: 542-488

- 542-520: First fossils of almost all phyla of animals

Ordovician period: 488-444

- 460-430: Andean-Saharan Ice Age
- 450: First known liverwort fossils

Silurian period: 444-416

- 425: First known vascular plant fossils
- 420: Likely origin of jawed fishes

Devonian period: 416-359

400: First known insect fossils

380-365: Origin of tetrapods

Carboniferous period: 359-299

360-260: Karoo Ice Age

314: First definite amniote fossils

300-200: Supercontinent Pangea

Permian period: 299-251

265: Origin of therapsids, mammal-like reptiles

251: End-Permian mass extinction

MESOZOIC era: 251-65

Triassic period: 251-202

230: Origin of dinosaurs

Jurassic period: 202-146

200: Break-up of Pangea into Gondwana and Laurasia

170: Diversification of mammaliforms and early mammals

150: Oldest known bird fossils

Cretaceous period: 146-65

125: First known fossils of flowering plants

125: First known fossils of placental mammals

65: KT mass extinction

CENOZOIC era: 65-present

~6: Divergence of *Pan* and *Homo* lineages

2.58-present: Pliocene-Quaternary Ice Age

History and perspectives

Among early developments in the Earth Sciences are a remarkably accurate estimate of the circumference of the Earth by Eratosthenes and an insightful description

of superposition of strata of sediments by Ibn Sina. Italian artist and scientist Leonardo da Vinci proposed in the XV century that fossils may be remains of long-dead animals. Principles of stratigraphy were formulated by a Danish scientist and Catholic bishop Nicolas Steno in the XVII century. Paleontology was established as a science, among others, by a French naturalist Georges Cuvier in late XVIII century, who was an advocate of catastrophism, a notion that in the past the Earth was affected by multiple drastic events that led to extinctions. A little later, an important book "Principles of Geology", published in 1830 by Charles Lyell, a friend of Darwin, established antiquity of the Earth and developed a concept of uniformitarianism. Of course, uniformitarianism, in contrast to gradualism, is perfectly consistent with catastrophism, but this was not obvious for quite a while. Quasiperiodic changes of the Earth's orbit were elucidated by a Serbian engineer Milutin Milanković early in the XX century. Radiometric dating was invented by a New Zealand chemist and physicist Ernest Rutherford in 1905. Plate tectonics stems from the ideas of a German geophysicist Alfred Wegener, proposed in 1912, but became universally accepted only 50 years later.

Modern techniques led to fascinating progress in Earth sciences, including more precise models of the structure of the Earth, better understanding of the mechanisms of plate tectonics, a variety of precise methods of radiometric dating, analysis of ratios of stable isotopes and other data on paleoenvironments, etc. Among these sciences, paleontology perhaps experienced the most astonishing progress in the XX century and, remarkably, this progress was more due to old-fashioned discovery of fossils than to anything else. Early Archean and Proterozoic fossils, including that of Ediacaran biota, as well as a number of transitional fossils shedding light on the origin of insects, tetrapods, mammals, birds, and humans greatly expanded the knowledge of the history of life.

Perspectives. Better methods and data will shed more light on the abiotic conditions experienced by past life. In particular, we will soon know much more about changes of the global environment in the Proterozoic eon, and the causes and details of the course of drastic past events that led to mass extinctions will be understood much better than today. We can also expect many important paleontological findings in the decades to come. These findings will continue to close gaps in our direct data on many major evolutionary transitions, at least those that occurred in Phanerozoic eon.