

Cenancestor

X Xia, University of Ottawa, Ottawa, ON, Canada

Q Yang, Nanjing Institute of Geology and Palaeontology, Chinese Academy of Science, Nanjing, China

© 2013 Elsevier Inc. All rights reserved.

This article is a revision of the previous edition article by W Fitch, volume 1, p 318, © 2001, Elsevier Inc.

Glossary

Dating A statistical protocol for inferring geological time of ancestral nodes on a phylogenetic tree.

Genetic code A specification of how triplet codons are translated into amino acids.

Homology Coancestry or sharing of a common ancestor, often inferred rather than proven. The maternal inheritance of mammalian mitochondria allows us to

infer that all human mitochondrial genomes are homologous. Sequence alignment helps us infer homologous sites in a set of nucleotide or amino acid sequences.

Horizontal gene transfer Acquisition of genetic materials of an organism, not from its immediate parent.

Molecular clock Constancy of substitution rates over time among different lineages.

Cenancestor is a term for the most recent universal common ancestor of all living organisms, where the Greek prefix 'cen' means recent (as in Cenozoic Era) and common or shared (as in coenocyte). It is synonymous to the last universal common ancestor (LUCA). Cenancestor is not necessarily the earliest form of life. It could have lived with many other organisms, except that all other organisms failed to leave any descendants in today's biosphere.

Contrary to the claim that universal common ancestry is a central pillar or the most fundamental premise of modern evolutionary theory, Darwinian evolutionary theory, either in its original or in its modern form, does not require the existence of a cenancestor. Finding life in another planet that does not even use DNA as genetic material does not hurt Darwinian evolutionary theory at all, even if that form of life comes to dwell among us and, in so doing, gives us indisputable evidence against universal common ancestry.

A recent statistical proof of the cenancestry hypothesis in 'Nature', however, represents a classical case of tautology. A set of highly conserved proteins sharing substantial sequence similarity were used to test whether they support the single-origin or multiple-origin hypothesis, and the single-origin hypothesis was strongly favored statistically. However, the test, as was conducted, can never fail to support the single-origin hypothesis given the sequence similarity. It is tautological because it assumes not only that homology leads to sequence similarity, but also that sequence similarity equals homology. The test in fact would even equate convergence to homology as unrelated genes that have presumably evolved in response to the same translation machinery also support the single-origin hypothesis.

There is overwhelming evidence that all living cellular forms of life, represented by Archaea, Eubacteria, and Eukarya, are genetically related. First, the genetic code is almost universally conserved. Second, different genes from representatives of these three kingdoms of life often yield similar phylogenetic relationships. Third, all cellular forms of life share substantial similarity in the three most fundamental biological processes (genome duplication, transcription, and translation). Fourth, the

common ancestry has been substantiated by numerous transitional fossils. Fifth, evolutionary convergence observed either at the phenotypic or at the DNA level, provides no sufficient explanation for the phenotypic or genetic similarities among organisms. However, whether living organisms share a cenancestor and what form the cenancestor takes remain controversial.

The earliest concept of a cenancestor is a cell from which all cellular forms of life are derived. If the cellular structure originated only once, then the existence of a cellular cenancestor is a logical necessity given the cell theory developed by Theodor Schwann, Matthias Jacob Schneider, and Rudolf Virchow in the late seventeenth century, which states that new cells are created by old cells dividing into two. However, it is difficult to demonstrate that the cellular structure originated only once. The concept also does not lead to any feasible way to build a universal phylogenetic tree and to identify the root of the tree (which is an operational definition of cenancestor).

The evolutionarily operational concept of cenancestor is a genome that codes a minimal set of core genes essential for cellular life (the minimal genome) and from which all other genomes are derived. Such a concept would theoretically allow us to find a shared set of core genes, build a universal tree based on the core genes, and to identify the root of the tree. Unfortunately, the shared set of core genes represents wistful thinking because a cellular function can often be performed by nonhomologous genes. Suppose that survival as a cellular form requires the cell to perform three essential functions F_1 , F_2 , and F_3 . If these three functions are performed by genes X_1 , X_2 , and X_3 in lineage 1, but by nonhomologous genes Y_1 , Y_2 , and Y_3 in lineage 2, then our set of core genes is empty. In general, the shared set of genes decreases toward zero as we increase the number of divergent genomes in the universal tree.

If we do find a shared set of core genes, albeit a small set, and if we do agree upon the tree of life hypothesis, which is controversial, how can we identify and date the root? Theoretically, the root can be identified if (1) the substitution process is non-stationary, for example, one daughter lineage becoming more

and more AT-rich and the other more and more GC-rich, then their parental node can be identified, or (2) the evolution of all lineages conform to a molecular clock. There are two forms of global molecular clock that are often not distinguished. The stringent form assumes a constant evolutionary rate among all lineages, and the relaxed form assumes synchronous evolution among all lineages (e.g., all lineages evolved fast during the Triassic, but slowly in the Jurassic, and fast again in the Cretaceous). The stringent clock is implicitly assumed in any statement of the global clock whereas the relaxed clock is implicitly assumed when statistical tests are conducted to test the validity of the global clock, that is, whether all leaves are equidistant from the root. While only a relaxed clock is required for identification of the root of a universal tree, dating the root requires the stringent global clock which is unlikely to hold among all lineages over a vast span of geological time. Even if such a stringent global clock is in operation, the substitution saturation would be so severe that deep phylogenetic information is lost. This would be especially true for those neutrally evolving sites that conform better to the molecular clock than other functionally constrained sites. Furthermore, the rampant occurrence of horizontal gene transfer suggests that the cenancestor is neither a single cell nor a single genome, but is instead an entangle bank of heterogeneous genomes with relatively free flow of genetic information. Out of this entangled bank of frolicking genomes arose probably many evolutionary lineages with gradually reduced rate of horizontal gene transfer confined mainly within individual lineages. Only three (Archaea, Eubacteria, and Eukarya) of these early lineages have representatives survived to this day.

See also: Genetic Code; Horizontal Gene Transfer; Molecular Clock; Phylogeny.

Further Reading

- Baptiste E, O'Malley MA, Beiko RG, *et al.* (2009) Prokaryotic evolution and the tree of life are two different things. *Biology Direct* 4: 34.
- Doolittle WF and Baptiste E (2007) Pattern pluralism and the Tree of Life hypothesis. *Proceedings of the National Academy of Sciences of the United States of America* 104: 2043–2049.
- Huelsenbeck JP, Bollback JP, and Levine AM (2002) Inferring the root of a phylogenetic tree. *Systematic Biology* 51: 32–43.
- Koonin EV and Wolf YI (2010) The common ancestry of life. *Biology Direct* 5: 64.
- McInerney JO, Pisani D, Baptiste E, and O'Connell MJ (2011) The public goods hypothesis for the evolution of life on Earth. *Biology Direct* 6: 41.
- Sober E (2008) *Evidence and Evolution: The Logic Behind the Science*. Cambridge: Cambridge University Press.
- Theobald DL (2010) A formal test of the theory of universal common ancestry. *Nature* 465: 219–222.
- Xia X and Lemey P (2009) Assessing substitution saturation with DAMBE. In: Lemey P, Salemi M, and Vandamme AM (eds.) *The Phylogenetic Handbook*, pp. 611–626. Cambridge: Cambridge University Press.
- Yonezawa T and Hasegawa M (2010) Was the universal common ancestry proved? *Nature* 468: E9.

Relevant Websites

- <http://www.ncbi.nlm.nih.gov/taxonomy> – Taxonomy.
- <http://www.timetree.org> – TimeTree.
- <http://www.treebase.org/treebase-web/home.html> – TreeBASE.
- <http://tolweb.org/tree> – Tree of Life.