

Molecules and the fossil record

If the fossil record cannot provide definitive answers to evolutionary questions perhaps we should be looking elsewhere. *Dr Blair Hedges* proposes that investigating gene sequences can reveal what fossils cannot.

BLAIR HEDGES

The fossil record contains much useful information on phylogeny and the times of species divergence but it is highly biased¹. These biases may be caused by differences in the abundance, habitat, or geographical range of a species, sparse sampling² or other factors. A case in point involves humans and our closest living relatives, the chimpanzees.

Humans and chimpanzees share a common ancestor which lived about 5-6 million years ago, but only fossils for the human lineage are known, providing many different hominid species³. The virtual lack of any fossil chimpanzees is most likely because chimps have lived in habitats - humid forests - where fossilization is rare.

Another problem is that the number of characters known for fossil species are often limited. Many extinct species are named from teeth, jaw fragments, or other small remnants. This poses problems in distinguishing one fossil species from another, and in trying to determine relationships using a limited number of characters. As a palaeontologist colleague of mine puts it, "fossils don't come with labels." They must first be identified before they become a useful part of the fossil record.

Fortunately molecules can give us an additional perspective on evolutionary history beyond what the fossil record can provide. Each living species has thousands of genes and millions of nucleotide sites that can be sampled for evolutionary studies.

Molecular clocks

Unlike morphology, where natural selection may obscure evolutionary history, sequences do not show such adaptive convergence⁴. Molecular data are thus more numerous and more useful for reconstructing phylogenies of living taxa. Molecular clocks, where differences in the sequences of genes are used to estimate the length of time since a common ancestor, also have an advantage over fossil clocks because they start counting mutations at the actual time of divergence. Fossil-based estimates of divergence time, on the other hand, are minimum estimates only.

Molecular clocks have revealed what appear to be major gaps in the fossil record of animals. Metazoan phyla apparently originated several hundred million years earlier than recorded by the Cambrian explosion of fossils⁵⁻⁷. Also, most orders of mammals appear suddenly in the early Cenozoic fossil record (~60 million years ago, Mya), yet molecular clocks record their diversification much earlier (~80-110 Mya)^{8,9}.

Molecular and morphological mismatches

Besides these timing differences, the relationships within the major group of mammals (those with a placenta) determined by molecules differs greatly from trees based on morphological data¹⁰. For example, evidence from several independent genes indicates that one-third of the living orders now form a well-supported group: the Afrotheria^{11,12}. This group includes such diverse forms as elephants, elephant shrews, tenrecs, golden moles, hyracoids,

sirenians and aardvarks. The group also makes geographic sense; the members are of African origin. Fossils and morphology, on the other hand, place elephants, sirenians, and hyracoids with ungulates (horses, cows, etc.)¹⁰.

The reaction from paleontologists and morphologists to these new discoveries has been mixed. Some have recognized the biases of the fossil record and have welcomed input from the molecular realm while others have not. In referring to this controversy over the origin of mammals, one paleontologist was quoted in *Science* recently as saying "if DNA clocks can't agree with the fossils then the problem is with the molecular clock."¹³

In this case, the researcher argued that there was no sampling bias involved. But this is only one of many potential biases in the fossil record¹. If mammals were less abundant before the K-T extinction of the dinosaurs, one should expect them to be less common in the Cretaceous fossil record. Any number of other biases may also be involved. The history of palaeontology would suggest that ruling out the possibility of earlier fossils (e.g., Cretaceous parrots¹⁴) is not a winning strategy.

Nothing is perfect

Molecular clocks and phylogenies are not without their limitations. First, there are known biases in the rates and types of substitutions that should be considered in using molecular data. Until now the biggest limitation has been the number of sequences available.

The last few years have seen some controversial proposals made by molecular phylogeneticists based on single genes or the mitochondrial genome (in one sense a single gene, albeit a large one). Some or all of these proposals may be correct, but evolutionary trees and divergence times based on only one or a few genes can be unreliable. It is even possible to obtain high confidence values for the wrong phylogeny if some biases are not taken into account¹⁵. Calibrations are an important issue with molecular clocks. If the calibration dates are not robust, then resulting time estimates will be of little use.

The way ahead

To understand the full impact molecules will have on reconstructing evolutionary history, one should look to the near future when large numbers of gene sequences will be available for many taxonomic groups. It is almost certain that this will lead to robust phylogenies and estimates of divergence time among living taxa. The fossil record will be of limited use in such cases, but will continue to be of immense value in reconstructing the evolutionary history of extinct taxa. The fossil record will also continue to give us insights into the morphology and ecology of past life.

Molecules and fossils, with their different strengths and weaknesses, are best viewed not as competing forces or equal characters to be blended but rather as complementary sources of information for evolutionary history.

S. Blair Hedges

Department of Biology, Pennsylvania State University, USA

References

1. Smith, A.B. *Systematics and the fossil record* (Blackwell Scientific Publications, London, 1994).
2. Martin, R.D. Primate origins: plugging the gaps. *Nature* **363**, 223-234 (1993).
3. Wood, B. The oldest hominid yet. *Nature* **371**, 280-281 (1994).
4. Doolittle, R.F. Convergent evolution: the need to be explicit. *TIBS*, **19**, 15-18 (1994).
5. Wray, G.A., Levinton, J.S. & Shapiro, L.H. Molecular evidence for deep precambrian divergences. *Science* **274**, 568-573 (1996).
6. Bromham, L., Rambaut, A., Fortey, R., Cooper, A. & Penny, D.

Testing the Cambrian explosion hypothesis by using a molecular dating technique. *Proc. Natl Acad. Sci. USA* **95**, 12386-12389 (1998).

7. Wang, D.Y.-C., Kumar, S. & Hedges, S.B. Divergence time estimates for the early history of animal phyla and the origin of plants, animals, and fungi. *Proc. R. Soc. Lond. B.* **in the press** (1998).
8. Hedges, S.B., Parker, P.H., Sibley, C.G. & Kumar, S. Continental breakup and the ordinal diversification of birds and mammals. *Nature* **381**, 226-229 (1996).
9. Kumar, S. & Hedges, S.B. A molecular timescale for vertebrate evolution. *Nature* **392**, 917-920 (1998).
10. Novacek, M.J. Mammalian phylogeny: shaking the tree. *Nature* **356**, 121-125 (1992).
11. Springer, M.S., *et al.* Endemic African mammals shake the phylogenetic tree. *Nature* **388**, 61-63 (1997).
12. Stanhope, M.J., *et al.* Molecular evidence for multiple origins of Insectivora and for a new order of endemic African insectivore mammals. *Proc. Natl Acad. Sci. USA* **95**, 9967-9972 (1998).
13. Normile, D. New views of the origins of mammals. *Science* **281**, 774-775 (1998).
14. Brutsaert, W. & Parlange, M.B. A lower jaw from a Cretaceous parrot. *Nature* **396**, 29-30 (1998).
15. Naylor, G.J.P. & Brown, W.M. Structural biology and phylogenetic estimation. *Nature* **388**, 527-528 (1997).