

Sequencing Proteins from Mastodon and T-Rex: Insights and Controversies

By Fernando Racimo

Old school paleontologists can tell a lot from fossil bones. A fossilized femur may help them infer how fast an animal could run from its predators. A fossilized mandible might give rough insights into the diet of a long-gone 6-ton carnivore like the T-Rex. But the macroscopic features of a fossil can only reveal so much about the beasts that roamed our planet in the past. To dig deeper, paleobiologists are starting to study microscopic organic compounds preserved through millions of years in remains of extinct organisms, expanding the nascent field of molecular paleontology.

Arguably the most popular of such compounds are two collagen fragments extracted from fossil remains of the *Tyrannosaurus rex* (1) and the *Mammuth americanus* or mastodon. The protein sequence of these macromolecules was analyzed using mass spectrometry (2). Scientists obtained sub-picomolar (extremely small) quantities of proteins from 68-million-year-old *Tyrannosaurus* soft tissue and a 600,000-year-old mammoth skull. The samples were fragmented into charged ions and sorted by mass. The resulting “fragmentation pattern” was compared to protein sequences of living mammals and reptiles to infer the actual peptide sequence of the proteins.

This is allegedly the first time a peptide sequence has been retrieved from fossilized material. But what can scientists do with a prehistoric string of amino acids? Researchers from Harvard University and North Carolina State University had an interesting answer: build a family tree. Phylogenetics is the study of how species are related, using similarities between them to build “phylogenetic trees” that reflect that relatedness. For example, birds are thought to be descended from dinosaurs because of morphological similarities between the two groups. Using a variety of statistical approaches, paleobiologists provided further support to this hypothesis: the T-Rex peptide sequences are more similar to sequences extracted from ostriches than to

alligator sequences. Similarly, the mastodon collagen sequences are more similar to elephant sequences than to the sequences of other extant mammals. (3).

There are, however, objections to using mass spectrometry on peptides that are thousands or millions of years old. Some scientists argue that the collagen remains could have undergone contamination and degradation and that more data are required to assess their validity (4), and others question the statistical significance of the sequences (5). The T-Rex samples in particular were buried in the Hell Creek Formation where temperatures can reach more than 20°C. Under those conditions, the half-life of collagen is approximately 2,000 years, much less time than the 68 million years that passed since the remains were first deposited. Nevertheless, the paleobiologists who carried out the sequencing have asserted that they validated their data through multiple proteomics tests and database searches. Moreover, the dinosaur tissue was found in exceptional conditions – buried under 1000 m³ of semiconsolidated sandstone – which might have protected them from environmental factors like water and air erosion (6).

The debate rages on but, if the data retrieved are indeed reliable, the potential for using mass spectrometry techniques in the retrieval of other ancient macromolecules is very promising. Protein sequences from extinct taxa could provide numerous insights into the way present and past species are related to each other and thus help uncover the secrets of life’s distant past. **H**

—*Fernando Racimo ‘11 is a sophomore concentrating in Organismic and Evolutionary Biology in Currier House*



▲ Left: Protein structure of a collagen fragment. Right: skull of *Tyrannosaurus Rex* and skeleton of *Mammuth Americanus*.

References

1. Schweitzer, M.H., et al., *Science* 316, 277-280 (2007).
2. Asara, J.M., et al. *Science* 316, 280 - 285 (2007).
3. Organ, C.L., et al., *Science* 320, 499. (2008).
4. Buckley, M., et al., *Science* 319, 33. (2008).
5. Pevzner, P.A., Kim, S., Ng, J., *Science* 321, 1040 (2008).
6. Asara, J.M, et al., *Science* 321, 1040 (2008).