Mass Spectrometry Analysis of Dinosaur Fossils Strengthens Evolutionary Linkages

Evolution is a controversial issue and many of today's hypotheses are largely based on the architecture of bones. With regards to the evolution of birds, most researchers have long suspected that they evolved from dinosaurs, but again this theory is based on the morphological similarities in bird and dinosaur skeletons. Historically, it was believed that the process of fossilisation destroyed any original protein material found inside bones. However, it has recently been discovered that fossilised bones from extinct taxa harbour the potential for obtaining protein or DNA sequences that could reveal evolutionary links to today's species. Identifying a relationship between the protein fragment sequences of dinosaurs and birds could allow scientists to positively conclude that the two species are evolutionarily related.

"The first challenge facing the research team was gathering enough protein to sequence"

Author Details:

John Asara PhD, Director,
Mass Spectrometry Core Facility,
Beth Israel Deaconess Medical Center
and Instructor in Pathology,
Harvard Medical School

Researchers from Harvard Medical School (HMS) and Beth Israel Deaconess Medical Centre (BIDMC) have been developing cutting-edge targeted proteomics methods for the precise sequencing of proteins in dinosaur fossils. The BIDMC is a teaching hospital of the HMS based in Boston. It features a Mass Spectrometry Core Facility, which provides the Harvard Medical Centre and other researchers with proteomics services. The majority of the laboratory's research efforts are specifically focused on developing proteomics-based strategies for sequencing miniscule amounts of protein in human tumours as well as ancient fossil proteomics. To conduct this research, BIDMC has developed mass spectrometry techniques using the Thermo Scientific LTQ linear ion trap and LTQ Orbitrap XL™ hybrid mass spectrometer. The researchers successfully captured and sequenced tiny pieces of collagen protein from a 68 million-year-old Tyrannosaurus rex (T. rex). The protein fragments appear to most closely match amino acid sequences found in collagen of present day chickens, leading support to the recent and still contentious proposal that birds and dinosaurs are evolutionary related.

THE CHALLENGES

The first challenge facing the research team was gathering enough protein to sequence. Tiny pieces of collagen protein were extracted from the bones of a 68 million year old T. rex and a 160,000 to 600,000 year old mastodon. Previous attempts to obtain protein or DNA sequences from such ancient fossils have failed because of extremely low concentrations of organic material remaining after extraction and because of degradation or modification of the remaining organic materials.

The scraps of dinosaur protein used in this application were wrested from a fossil femur discovered by scientists from the Museum of the Rockies in 2003 in Hell Creek Formation, a barren fossil-rich stretch of land that spans several states, including Wyoming and Montana. The bone extract arrived in the BIDMC laboratory in the form of a gritty brown powder that had to be rid of contaminants. After removing the minerals and impurities from the sample, less than a billionth of a gram of protein was left for analysis. The second challenge was to interpret the amino acid sequences. For extinct species, the real goal is to find sequences unique to that organism.

Mass spectrometry has been extensively used for protein identification. However, this traditional method does not provide sufficient sensitivity and duty cycles required in cases of complex proteins. For example, mass spectrometry has successfully been used to report protein sequences from young dinosaur fossils, but when the fossils are one million years old and above, the protein concentrations are so small that they fall below the limits of detection of most mass spectrometers. Analysis has also been hindered by theoretical limits based on predicted rates of degradation. In addition, most commercial software for identifying peptide sequences by mass spectrometry relies on the peptide fragmentation pattern matching identically to that of a peptide/protein sequence in existing sequence databases.

Researchers from BIDMC identified the clear need for a more sensitive and reliable two-step mass spectrometry approach, capable of increasing throughput and depth of analyses for the demanding identification of proteins in dinosaur fossils.

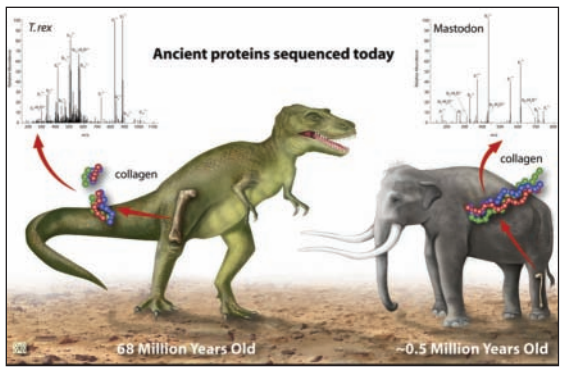

Figure 1. Researchers from Harvard Medical School and Beth Israel Deaconess Medical Center captured and sequenced tiny pieces of collagen protein from a 68 million-year-old Tyrannosaurus rex and a 160,000- to 600,000-year-old mastodon.

THE ANALYSES

Using the Thermo Scientific instruments, the BIDMC team was able to sequence tiny pieces of collagen protein. Investigators managed to purify the T. rex dinosaur fossil protein and break it down into fragments or peptides that were 10 to 20 amino acids long using the enzyme trypsin. The peptides were then purified and concentrated using micro-chromatography and passed over a liquid chromatography (LC) column, where they were separated from one another and then sprayed into the mass spectrometers at extremely low nanoliter flow rates for optimal sensitivity. In addition, Asara and his colleagues isolated and sequenced more than 70 protein fragments from the mastodon.

Typically, a mass spectrometer measures the mass-to-charge ratio of peptides as they come off the LC column. In order to maximise yield, Asara and his laboratory staff used a linear ion trap, which captures and holds peptides through time, delivering more structural information. The collected peptides were measured for mass and were subsequently isolated and fragmented to reveal their amino acid sequence. Using this two-step, or tandem (MS/MS) procedure, seven separate strings of amino acid were netted.

The next step involved the interpretation of the MS/MS data. The acquired data is compared to a database of theoretical MS/MS spectra of existing amino acid sequences. As collagen is a highly conserved protein, it was highly likely that some of the dinosaur peptide sequences matched those of an existing species.

THE FINDINGS

Of the six T. rex peptides identified, five were of a particular class of collagen protein, collagen alpha I. The majority of these peptides were found to be identical matches to amino acid sequences found in chicken collagen alpha I, lending support to the proposal that birds and dinosaurs are evolutionarily related. Other protein fragments matched those found in Japanese newts and frogs. Using molecular evidence, the BIDMC team has been able to place the T. rex (a non-avian dinosaur) on the animal kingdom's family phylogenetic tree that traces the evolution of species.

The mere existence of such exceedingly ancient protein defies a longstanding assumption. For centuries, it was believed that when an animal dies, protein immediately begins to degrade and, in the case of fossils, it is slowly replaced by mineral. This substitution process was thought to be complete after one million years. As a consequence, scientists were reluctant to examine really old bones. The breakthrough research conducted by the BIDMC and NCSU scientists has proven that it is actually possible to identify sequences far beyond one million years.

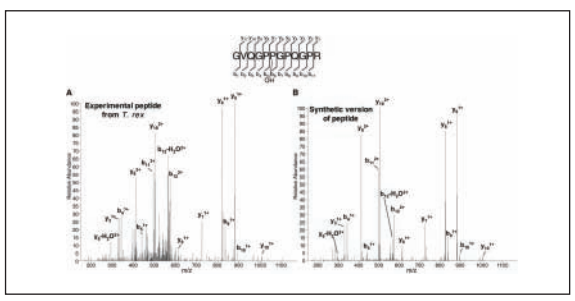

Figure 2. Fragmentation pattern obtained from analyzing the T. rex peptide with state-of-the-art mass spectrometry technology. Most of the protein (collagen) sequences from the T. rex sample uniquely match collagen sequences belonging to modern-day chickens that are found in publicly-available protein databases.

A set of theoretical collagen protein sequences were generated, that represented the types that might have been present around the time of T. rex. Unsurprisingly, none of the dinosaur peptides matched the theoretical set. Mastodon bone was tested against a database of existing amino acid sequences and against a set of mastodon theoretical sequences. Contrary to the T. rex findings, a total of 78 peptides were identified, including four unique sequences. As a result of the analysis, the researchers proposed a close phylogenetic relationship between the mastodon and modern elephants.

THE TECHNOLOGY

The advanced mass spectrometry methods played critical roles in achieving protein identification in BIDMC's fossilomics analyses. The linear ion trap method substantially improved sensitivity and scan speed increasing throughput and depth of analyses. Hybrid mass spectrometry provided high MS/MS mass accuracy and resolution measurements, in addition to acquiring very low molecular weight fragment ions and also enabled the team to sort out minor sequence discrepancies. With this technology, all forms of proteomics analysis are possible, from qualitative to quantitative through stable isotope labelling and label free methods; for example it has enabled stable isotope labelling experiments such as SILAC to be performed with unparalleled scan speed and sensitivity. When it was first implemented, some previous samples were re-analysed to find phosphorylation sites and more sites were identified using this technique.

Hybrid mass spectrometry is a powerful technique used for the challenging analysis of low-level components in complex mixtures. The method offered superior mass accuracy for lower false positive rates (FPRs) while its excellent MS/MS sensitivity and dynamic range led to more protein identifications.

An important aspect of proteomics is not only the identification of all proteins in complex biological samples but also the accurate determination of their relative concentrations. Hybrid mass spectrometry is ideally suited to perform such experiments.

DISCUSSION

The ability to sequence intact peptides from a 68 million year old source is attributed to several factors, including the exceptional preservation of the soft tissues from the Hell Creek environment, the fresh preparation of the fossil samples without curation or preservation and the advancements in the sensitivity of mass spectrometry technology over the past decade.

The fact that collagen was abundant in the mastodon sample, which could be approximately half a million years old, also sheds light on the fact that protein lasts much longer than one million years.

CONCLUSION

Obtaining genome sequences from a number of taxa has dramatically enhanced the ability to study the evolution and adaptation of organisms. The implementation of sophisticated mass spectrometry techniques at the BIDMC laboratory has expanded the facility's analytical capabilities and opened the door to cutting-edge fossilomics research. The wide dynamic range, improved sensitivity, fast cycle times, ultra high mass accuracy and superior resolution of mass spectrometry has enabled the development of advanced proteomics methods, allowing researchers to obtain complex protein sequences from comparably aged species and to identify some of these sequences as unique to those species.

This research has set the basis for acquiring the capability to clearly define the evolutionary relationship between extinct and current species. Mass spectrometry can be used to determine unique sequences from ancient organisms from peptide fragmentation patterns, a valuable tool to study the evolution and adaptation of ancient taxa from which genomic sequences are unlikely to be obtained.

Additionally, the research is projected to have significant implications for human health. The mass spectrometry and bioinformatics based approach can be applied not only to obtain sequences from extinct organisms, but also to obtain protein sequences from present day organisms whose genomes have not been sequenced and to discover mutations in diseased tissues such as cancers. It can be used to develop new protein-based tests for prostate cancer and other diseases. Scientists hope that the method will allow them to pick up mutations in the proteins that would infer what the DNA mutations were. This could lead to new tests for detecting cancer at the disease's earliest stages.