NATURE | NEWS

Mystery of Darwin's 'strange animals' solved

Analysis using ancient collagen protein could permit study of fossils older than DNA allows.

Jo Marchant

18 March 2015

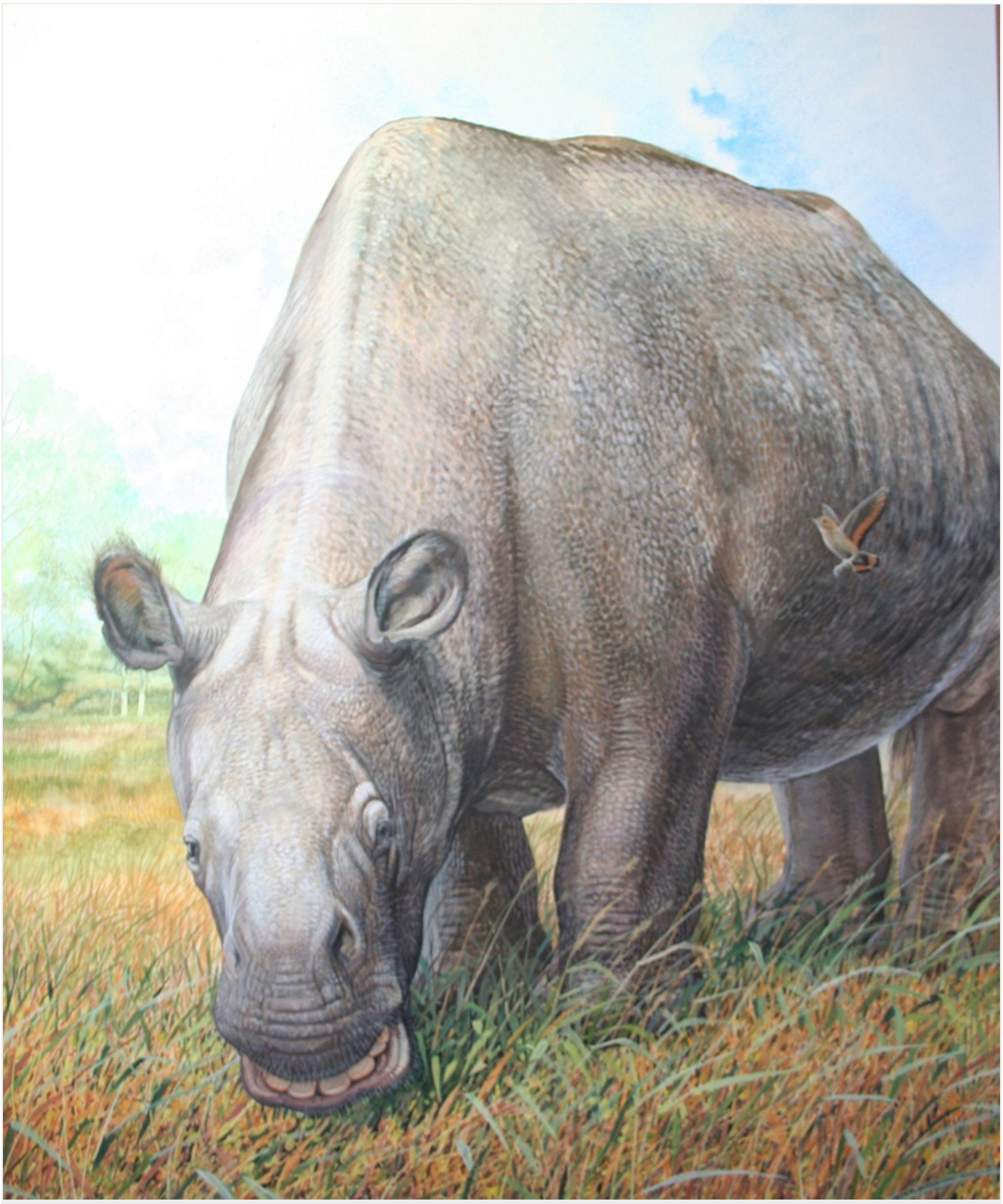

Illustration by Peter Schouten/"Biggest, Fiercest, Strangest"/W. Norton Publishers

An artist's impression of the *Toxodon* — body of a rhino, head of a hippo, teeth of a rodent. Collagen sequences from its fossil suggest that the animal is related to horses, tapirs and rhinos.

When Charles Darwin visited South America on HMS *Beagle* in the 1830s, he discovered fossils of several hefty mammals that defied classification, such as *Macrauchenia*, which looked like a humpless camel with a long snout; or *Toxodon*, with a rhino’s body, hippo’s head and rodent-like teeth — which he described as “perhaps one of the strangest animals ever discovered”.

Since Darwin's time, no-one has been able to work out where the bizarre beasts fit in the mammalian family tree[1]. But now, by analysing ancient collagen protein from 12,000-year-old fossils, researchers say that they have solved the puzzle. The scientists behind the work also think that ancient proteins could revolutionize the study of long-extinct species, revealing the secrets of fossils millions of years older than can be studied using DNA.

Part of a group of more than 250 mammals known as the South American ungulates, the creatures lived on the continent for around 60 million years before disappearing around 12,000 years ago. The confusion over their ancestry is partly due to a fragmentary fossil record, but also because researchers have had no luck isolating DNA from South American ungulate fossils; the molecule degrades quickly in the continent's warm climate.

So Ian Barnes, a molecular evolutionary biologist at the Natural History Museum in London, teamed up with bioarchaeologist Matthew Collins at the University of York, UK, and an international team of researchers to try a different tactic: extracting collagen. The protein survives around ten times as long as DNA and is a major structural component of bone. "Compared to DNA, there's absolutely tons of it," Barnes says.

The team first built a collagen family tree, which laid out the collagen sequences of different mammals on the basis of their familial relationships. The researchers had to extract and sequence collagen from tapirs, hippos and aardvarks to build up their picture. With that in hand, they sequenced collagen from four ungulate specimens from two different museums in Argentina — two *Toxodon* specimens around 12,000 years old and two *Macrauchenia* that could not be carbon-dated — and compared the ancient proteins against their tree.

South American ungulates were recently suggested — on the basis of their fossils — to be part of the group Afrotheria, along with elephants and manatees[1]. But the protein sequences, reported in this week's *Nature*, revealed that the museum specimens are most closely related to Perissodactyla, a group that includes horses, tapirs and rhinos[2].

The study is "a big step forward", says Rob Asher, a palaeobiologist at the University of Cambridge, UK, who was not part of the research team. Biologists can now start to tease out how the fossils' physical traits evolved, he says.

The ancient protein revolution

But the researchers who took part in the study say that this is just the start. Ancient proteins could now prove as revolutionary as DNA for studying the tree of life, "but with the possibility of reaching much further back in time", they write. Modern instruments mean that researchers are now getting many times more information from ancient samples, says bioarchaeologist Collins: today's mass spectrometers can measure protein masses more accurately than before, and can collect and concentrate rare components within protein mixtures.

The oldest DNA recovered so far, from an ice core in Greenland, is 450,000–800,000 years old[3]. Proteins could reach back millions of years. In 2007, US researchers claimed to have sequenced collagen fragments from a 68-million-year-old dinosaur fossil[4], but the result has proved controversial. Otherwise the oldest protein sequence recovered so far is from a camel that lived in the Arctic 3.2 million years ago[5]. "Certainly 4 million years will not be a problem," says Collins. "In cold places, maybe up to 20 million years."

Proteins could also be useful for studying extinct species that lived more recently in hotter environments where DNA studies are difficult: what Collins describes as "weird and wonderful" animals around during the late Pleistocene, from the dwarf elephants and enormous rodents of the Indonesian island of Flores to Australia's giant lizards and kangaroos.

Asher isn't convinced that ancient proteins will be as revolutionary as DNA, but he describes the potential as "very exciting". For species that have become extinct within the last few million years, he says, "this could really rock the boat".

Nature | doi:10.1038/nature.2015.17138

References

1. Agnolin, F. L. & Chimento, N. R. *Mamm. Biol.* **76**, 101–108 (2011).
2. Welker, F. *et al*. *Nature* http://dx.doi.org/10.1038/nature14249 (18 March 2015).
3. Willerslev, E. *et al*. *Science* **317**, 111–114 (2007).
4. Asara, J. M. *et al*. *Science* **316**, 280–285 (2007).
5. Rybczynski, N. *et al*. *Nature Commun.* **4**, 1550 (2013).