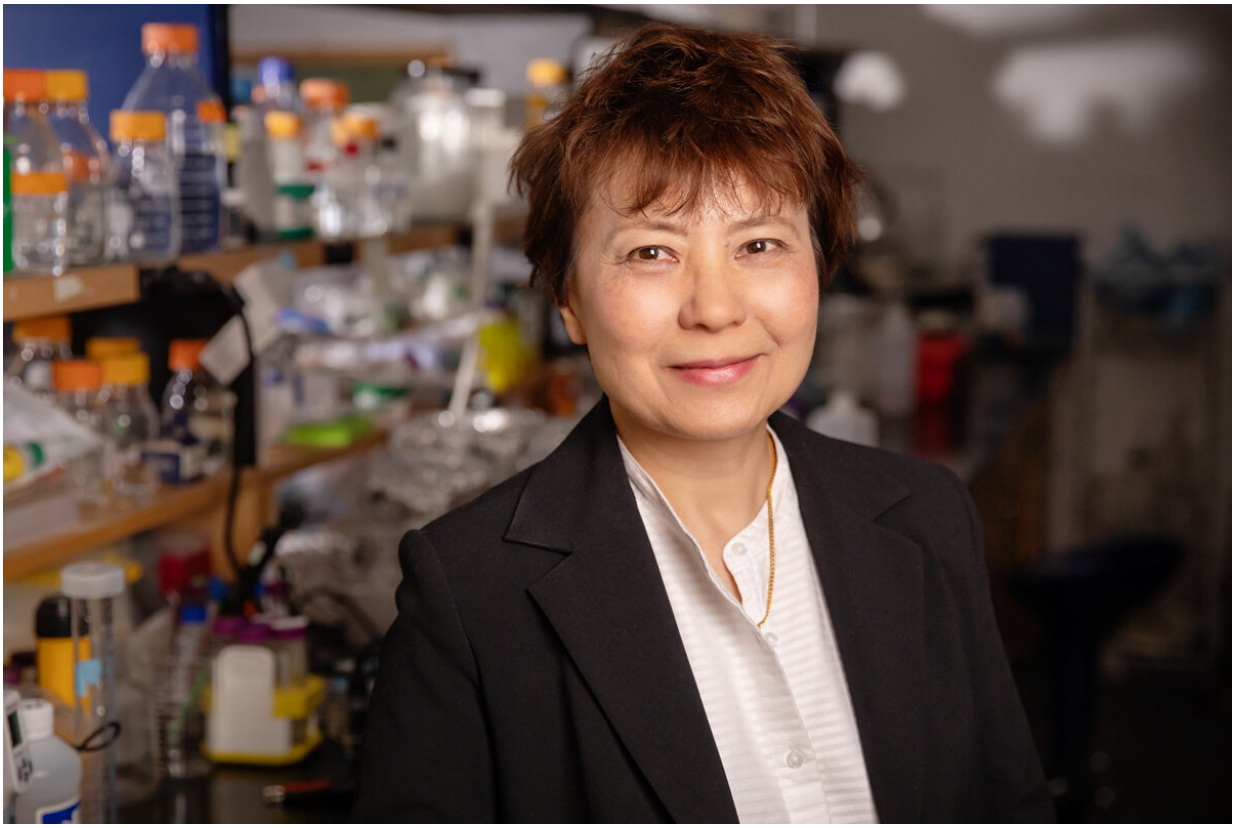
Study of Arctic fishes reveals the birth of a gene—from 'junk'

February 11 2019



University of Illinois animal biology professor Christina Cheng and her colleagues determined how the gene for an antifreeze protein in Arctic fish evolved from noncoding DNA. Credit: L. Brian Stauffer

Though separated by a world of ocean, and unrelated to each other, two

fish groups—one in the Arctic, the other in the Antarctic—share a surprising survival strategy: They both have evolved the ability to produce the same special brand of antifreeze protein in their tissues. A new study describes in molecular detail how the Arctic fishes built the gene for their antifreeze from tiny fragments of noncoding DNA, regions once considered "junk DNA."

The findings are reported in the *Proceedings of the National Academy of Sciences*.

"Years ago, we discovered how antifreeze glycoproteins evolved in Antarctic notothenioid fishes, and we knew that the Arctic cod evolved an identical version—but not in the same way," said University of Illinois animal biology professor Christina Cheng, who led the new study with graduate student Xuan Zhuang. "But exactly how the codfish independently did it has remained a lasting puzzle."

To solve that puzzle, Cheng and her colleagues scoured fish and other vertebrate genomes for a gene that might have been the ancestral precursor to the codfish antifreeze gene. They came up empty, so they decided to compare the genomes of codfish that did and did not produce antifreeze protein to see how the two lineages differed. The researchers found the ancestor of the antifreeze gene in a region of noncoding DNA, which, as its name implies, does not code for a viable protein.

"For many years after this discovery, I thought nobody was going to believe me, because the prevailing mindset at that time was that new genes have to evolve from pre-existing protein-coding gene ancestors," Cheng said.

Eventually, the researchers pieced together the details of how the codfish antifreeze gene originated.

"Its development in these fishes that make their living in icy Arctic waters occurred as a result of a series of seemingly improbable, serendipitous events," Cheng said.

Not just any random DNA sequence can produce a viable protein—let alone a lifesaving one like the antifreeze protein, Cheng said. Even if the original sequence contained the right order of building blocks that allow it to undergo transcription from DNA to RNA—the first step in building a protein—several hurdles remain. Specific sequences determine whether and how genes are transcribed into RNA, how they are edited and whether they are then translated from RNA into proteins.

In the case of a secreted protein like the antifreeze protein in codfish, a specific "signal sequence" also is required to process the final protein properly and maneuver it out of the cell and into the bloodstream.

The codfish antifreeze protein gene was assembled as a result of several molecular events, the study found. At its heart, a tiny segment of noncoding DNA, consisting of nine building blocks called nucleotides, underwent multiple duplications, creating a longer sequence of repeats. These code for a repeating series of three amino acids: threonine-alanine-alanine. These amino-acid repeats have just the right chemical properties to bind to ice crystals in the blood and prevent the crystals from growing.

Several other serendipitous events occurred in the evolving gene sequence, Zhuang said. One bit of DNA—when edited in just the right way after the gene is transcribed from DNA to RNA—included a sequence that tags the protein for export from the cell. A random one-nucleotide deletion shifted how the gene would be translated, linking the secretion signal to the region of antifreeze repeats, making them part of the same gene.

And, somehow, the gene also obtained the proper control sequence that

would allow the new gene to be transcribed into RNA. This transcription signal may have been inserted from elsewhere in the genome. Or, Zhuang said, the rest of the gene may have wandered from its original location to one that contained a transcription signal. Such DNA "translocation" events are a common occurrence across the genome.

The findings offer fresh insights into how a cell can invent "a new, functional gene from scratch," Cheng said.

"Evolution is not that efficient," she said. "It's a make-do kind of thing."

The cellular machinery appears to be constantly cranking out transcripts of DNA sequences that may or may not code for functional proteins.

"This process seems wasteful, but the cell can recycle the RNA that doesn't get used," Cheng said.

Only if the RNA transcript also contains certain other sequences will it be translated into protein, she said. If that protein happens to give the organism an advantage—ensuring its survival in icy Arctic waters, for example—the corresponding DNA sequence becomes "fixed" in the genome. Only those individuals that have that sequence in their genomes will persist in the environment and pass on the new trait.

"After years of study, we finally understand the birth of the codfish antifreeze gene," Cheng said. "This paper explains how the antifreeze protein in the northern codfish evolved. And it's an even more fascinating mechanism than the Antarctic version, which involved a pre-existing gene."

More information: Xuan Zhuang el al., "Molecular mechanism and history of non-sense to sense evolution of antifreeze glycoprotein gene in northern gadids," *PNAS* (2019).

www.pnas.org/cgi/doi/10.1073/pnas.1817138116