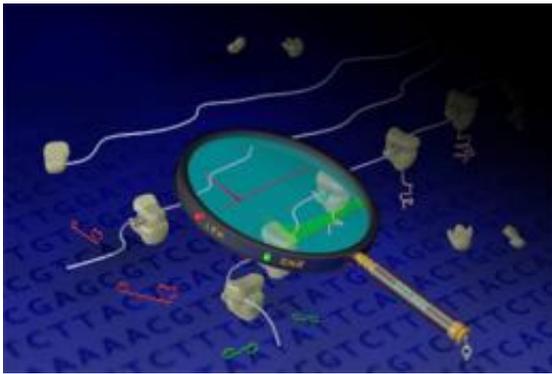


New method helps researchers decode genomes

August 28 2012, by Krishna Ramanujan



In cells, a messenger RNA (purple) is decoded by ribosomes (beige), which in turn produce the desired chains of amino acids that make up proteins, called polypeptides (pink coil). In order to understand where a gene's code begins, two related but distinct translation inhibitors (red and green) are used by the researchers to freeze the DNA translation process. Image: Shu-Bing Qian

(Phys.org)—Although scientists sequenced the entire human genome more than 10 years ago, much work remains to understand what proteins all those genes code for.

Now, a study published online Aug. 27 in the [Proceedings of the National Academy of Sciences](#) describes a new approach that allows researchers to decode the genome by understanding where genes begin to encode for polypeptides, long chains of amino acids that make up proteins.

"The key to decoding the genome is knowing exactly where the genes start to encode polypeptides," said Shu-Bing Qian, the paper's senior author and an assistant professor of [nutritional sciences](#) at Cornell. "If we know where they start, then we can predict what proteins they produce based on the gene's sequence."

[Gene sequences](#) are composed of four nucleotides—adenosine (A), cytidine (C), guanosine (G) and thymidine (T)—but the codes are arranged by three consecutive [nucleotides](#). The problem is, depending on where one begins to read the code, a single segment of DNA can generate different [gene products](#).

The new approach takes advantage of ribosomes, the translation machinery that decodes [messenger RNA \(mRNA\)](#), which carries the coding information from the DNA and translates those codes into chains of [amino acids](#), proteins' building blocks.

When translating mRNA, the ribosome at the start position has an empty space inside. Qian and colleagues used a special [chemical compound](#) that fills in that empty space and freezes that [ribosome](#). This allows the researchers to locate precisely where a gene starts to encode polypeptides. They then use that information to predict what proteins are produced from the sequence.

By using this method, the researchers found that the same mRNA can have multiple start sites that lead to production of different proteins.

"About 50 percent of mRNA has more than one start site," said Qian. In this way, a limited genome can have multiple possibilities, depending on where on the gene a start site occurs. For instance, if it occurs later in a gene's sequence, it can code for a shorter or totally different protein.

During transcription, mRNA substitutes uracil (U) for T found in DNA.

"Traditionally, all the known translation start sites were AUG. But we found that other codons, such as CUG can also serve as a start site," Qian said. The finding will rewrite the conventional thinking about genes and where they start to encode, he added.

The results suggest that the entire complement of proteins that can be expressed by a single gene is much more diverse than previously thought. Also, predicting what proteins a gene can code for may be much more challenging because of this alternative decoding process.

The technique can also be used to examine the genome of viruses, which are known for hijacking a cell's translation machinery to create new viruses.

"Viruses often use this alternative translation to maximize the coding capacity of their limited genome sequence to generate viral proteins," Qian said. This method has the potential to discover new viral proteins, he added.

Sooncheol Lee, a former postdoctoral researcher, and Botao Liu, a graduate student, both in Qian's lab, are the paper's lead authors, and Ben Shen, a chemist at Scripps Research Institute, was a co-author.

The study was funded by the National Institutes of Health, Ellison Medical Foundation and the U.S. Department of Defense.

Provided by Cornell University

Citation: New method helps researchers decode genomes (2012, August 28) retrieved 26 April 2024 from <https://phys.org/news/2012-08-method-decode-genomes.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private

study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.