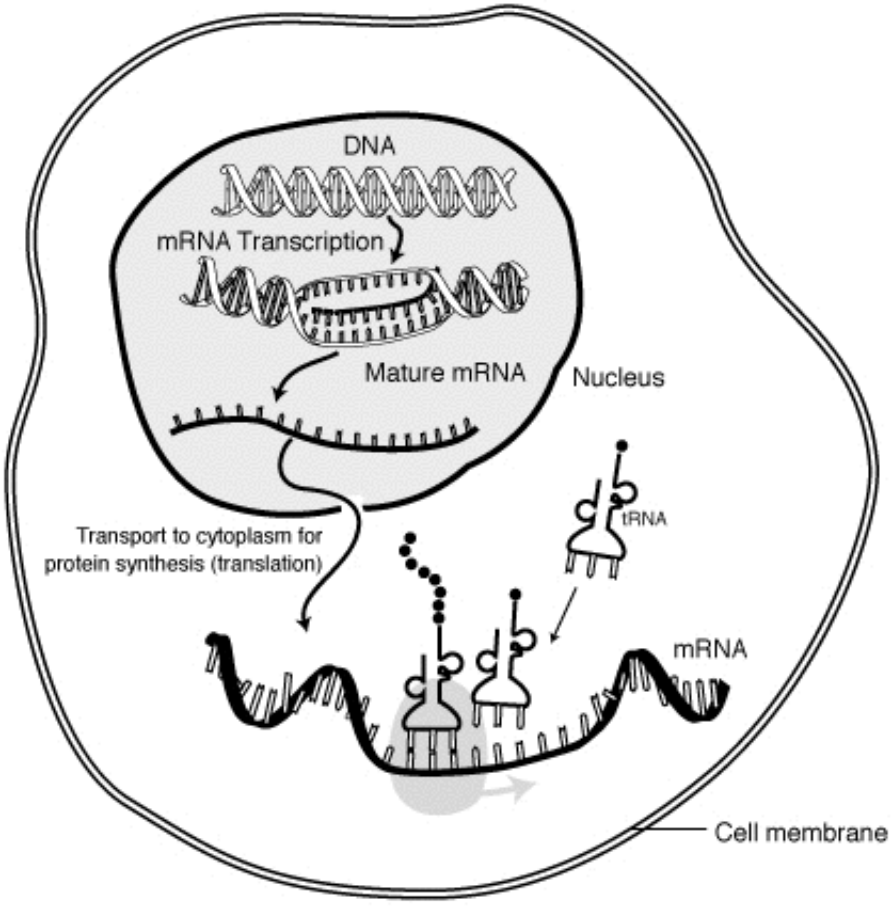
Researchers find hidden meaning and 'speed limits' within genetic code

March 12 2015



The "life cycle" of an mRNA in a eukaryotic cell. RNA is transcribed in the nucleus; processing, it is transported to the cytoplasm and translated by the ribosome. Finally, the mRNA is degraded. Credit: Public Domain

Case Western Reserve scientists have discovered that speed matters when it comes to how messenger RNA (mRNA) deciphers critical information within the genetic code—the complex chain of instructions critical to sustaining life. The investigators' findings, which appear in the March 12 journal *Cell*, give scientists critical new information in determining how best to engage cells to treat illness—and, ultimately, keep them from emerging in the first place.

"Our discovery is that the genetic code is more complex than we knew," said senior researcher Jeff Coller, PhD, associate professor, Division of General Medical Sciences, and associate director, The Center for RNA Molecular Biology, Case Western Reserve University School of Medicine. "With this information, researchers can manipulate the genetic code to achieve more predictable outcomes in an exquisite fashion."

The genetic code is a system of instructions embedded within DNA. The code tells a cell how to generate proteins that control cellular functions. mRNA transmits the instructions from DNA to ribosomes. Ribosomes translate the information contained within the mRNA and produce the instructed protein. The genetic code comprises 61 words, called "codons," and a single codon, a sequence of three nucleotides, instructs the ribosome how to build proteins.

The code not only dictates what amino acids are incorporated into proteins, it also tells the cell how fast they should be incorporated. With this information, researchers can manipulate the genetic code to achieve predictable protein levels in an exquisite fashion."

The most significant breakthrough in the Case Western Reserve work is that all of the words, or codons, in the genetic code are deciphered at different rates; some are deciphered rapidly while others are deciphered slowly. The speed of how mRNA decodes its information is the sum of

all the codons it contains. This imposed speed limit then ultimately affects the amount of protein produced. Sometimes faster is better to express a high level of protein. Sometimes slower is better to limit the amount protein. Importantly, codons are redundant—many of these words mean the same thing.

Coller and colleagues found that each of the codons is recognized differently by a ribosome. Some codons are recognized faster than others, but these differences in speed are tiny. Over the entire span of an mRNA, however, each tiny difference in speed is powerfully additive.

"Many codons mean the same thing, but they influence decoding rate differently. Because of this, we can change an mRNA without changing its protein sequence and cause it to be highly expressed or poorly expressed and anywhere in between," he said. "We can literally dial up or down protein levels any way we want now that we know this information."

During their research, investigators measured the mRNA decay rate for every transcript within the cell. They were seeking answers for why different RNAs had different stabilities. With statistical analysis, investigators compared the half-lives of mRNAs to the codons used within these messages. A strong correlation emerged between codon identity and mRNA message stability. They ultimately linked these observations back to the process of mRNA translation.

"mRNA translation and mRNA decay are intimately connected. This can be very beneficial to scientists. If you would like a gene to be expressed really well, you simply change the protein sequence to be derived by all optimal codons. This will both stabilize the mRNA and cause it to be translated more efficiently," Coller said. "If you need an mRNA to express at a low level, you fill it with non-optimal codons. The mRNA will be poorly translated and thus unstable. Evolution has used codon

optimization to shape the expression of the proteome. Genes of similar function use similar codons; therefore, they are expressed at similar levels."

His discovery has a variety of practical implications for medicine. From a bioengineering perspective, molecular biology techniques can be applied to manipulate the gene to contain ideal codons and obtain the gene expression pattern that is most beneficial to the application. From a human physiological standpoint, it's possible to learn the speed limit for each and every mRNA and then determine if this changes in specific pathologies such as cancer. Currently, it is unknown whether codons convey different speeds in disease states. A future direction for research will be to link codon speeds to specific illnesses. The potential is also there to develop drugs that can manipulate higher or lower gene expression by changing the decoding rate.

Codon activity also may also provide important clues about the source of many illnesses that have not been linked to specific gene mutations. Altering codon-dependent translation rates has the potential to change protein function profoundly, and no primary mutation will be detected. Rather the problem is not the gene itself, but the factors that influence decoding rates. Codon-dependent speed limits may underlie the cause of whole classes of disease states. For example, a recent study suggests that in more than 450 different cancer samples, factors influencing codon-dependent speed limits might be changing.

"The sky is the limit," Coller said. "Since this finding is so new, we have no idea what the potential is. The next step is to determine if changes in decoding speed can be an underlying mechanism that alters gene expression in human disease."

Provided by Case Western Reserve University