Science Education Collection

An Overview of Gene Expression

URL: https://www.jove.com/science-education/5546

Abstract

Gene expression is the complex process where a cell uses its genetic information to make functional products. This process is regulated at multiple stages, and any misregulation could lead to diseases such as cancer.

This video highlights important historical discoveries relating to gene expression, including the understanding of how distinct combinations of DNA bases encode the amino acids that make up proteins. Key questions in the field of gene expression research are explored, followed by a discussion of several techniques used to measure gene expression and investigate its regulation. Finally, we look at how scientists are currently using these techniques to study gene expression.

Transcript

Gene expression is the process where information contained in a cell's DNA is used to make functional products. This carefully orchestrated procedure is regulated at several stages, and misregulation can often result in diseases such as cancer.

This video provides an overview of the history of gene expression research, key questions, methods in the field, and how these techniques are being applied.

We'll begin by reviewing some major discoveries about gene expression.

The first convincing model for how DNA might carry genetic information was established in 1953, when Francis Crick and James Watson, with help from Rosalind Franklin's data, solved the structure of DNA—a double helix made of two linear chains of nucleotide bases that are arranged in a defined, but infinitely variable, sequence.

Five years later, Crick proposed two important ideas that would form the backbone of our understanding of gene expression. His "sequence hypothesis" suggested that DNA's nucleotide sequence is used, via an unstable RNA intermediate, as a code for proteins' amino acid sequences. At the same time, his "central dogma" hypothesized the different flows of genetic information that can occur, and in particular, held that information cannot be transferred from protein back to nucleic acids.

In 1960, François Jacob and Jacques Monod—through their work on lactose-metabolizing genes in bacteria—proposed a model for the regulation of gene expression. They suggested that the expression of "structural genes," which perform structural or enzymatic functions, is controlled by the products of "regulatory genes" that bind to adjacent regulatory sites. We now know that substantially similar modes of gene regulation, mediated by proteins called transcription factors, occur in all organisms.

A year later, in 1961, Jacob—along with Sydney Brenner—discovered messenger or "m"-RNA as the unstable intermediate between DNA and proteins proposed by Crick. That same year, Brenner and Crick began to crack the "genetic code," which dictates how information in DNA encodes proteins. They determined that each triplet of adjacent nucleotides, or a "codon," specifies one of the 20 amino acids that constitute proteins.

Over the next few years, researchers led by Marshall Nirenberg, Har Gobind Khorana, and Severo Ochoa used multiple approaches to define the amino acids encoded by all 64 possible codons. With the cracking of the genetic code, scientists continued to investigate how gene expression is regulated.

A major discovery came in 1974, when Roger Kornberg and colleagues showed that DNA in eukaryotic cells, such as those of animals and plants, is "wrapped" around complexes of histone proteins, yielding structures now called "nucleosomes." We now know that changes in chromatin structure play important roles in gene regulation.

Another twist came in 1977, when Phil Sharp and Rich Roberts found that mRNA sequences were not entirely complementary to their corresponding DNA templates. Certain "missing regions," now called introns, are removed from between the protein-coding exons in the mature RNA transcript in a process known as "splicing." Five years later, the research group of Ronald Evans demonstrated that "alternative" splicing of the same transcript could produce variants, or "isoforms," of the same protein with different functions.

Since the 1990s, our understanding of the complexity of gene regulatory networks expanded dramatically with the discovery of RNA-mediated gene silencing. We now know that several different families of small RNAs, with members that are 20−30 nucleotides in size, regulate gene expression in a variety of ways.

After reviewing the history of gene expression research, let's look at some major questions in the field.

One topic being investigated is how transcription factors regulate genes. Scientists are not only interested in identifying the genomic sequences bound by transcription factors, but are also looking at how regulatory proteins interact with one another to integrate signals and regulate gene expression.

Other researchers study alternative splicing, and how this process is regulated in different biological contexts. In addition, some of them are trying to determine whether protein isoforms always have different, distinct functions.

Finally, many researchers are investigating the mechanism of action of small RNAs, and are trying to identify their regulatory targets. There is also a growing interest in whether small RNAs can be used as "biomarkers" to diagnose diseases.

Now, let's look at the tools researchers use to assess gene expression.

A popular method is reverse transcription, or "RT"-PCR, which converts RNA into complementary or "c"-DNA before subjecting it to amplification. By including fluorescent molecules that are incorporated into the DNA during PCR, it's possible to use this technique to quantitatively measure gene expression and observe the results in "real-time."

To simultaneously assess the expression of thousands of genes, microarrays can be used. Here, DNA sequences are "printed" on slides, which are then hybridized to fluorescent probes generated from sample RNA. The resulting pattern of fluorescence can be used to identify expressed genes.

Another technique to profile gene expression is to sequence the transcriptome, or all of the expressed RNAs in a cell. Here, cDNA generated from RNA samples is subjected to high-throughput sequencing. Unlike microarrays, transcriptome sequencing does not require preexisting genomic information, and can be used to identify unknown transcripts or novel gene isoforms.

Researchers can also visually evaluate where a gene is expressed using *in situ* hybridization. In this technique, RNA is first hybridized with complementary probes, which can then be recognized by enzyme-conjugated antibodies that produce a visual color or fluorescence signal.

The reporter assay is another technique that can provide insight into gene regulation. The reporter gene product generates a signal such as color or fluorescence. The reporter maybe fused directly to a gene of interest, or be placed under the control of a regulatory sequence, such as the promoter that drives a gene's transcription, or a more distant enhancer element. The reporter signal can then act as readout for the regulatory element's activity, or the expression pattern of the gene of interest.

Finally, chromatin immunoprecipitation or "ChIP" can be used to identify the genomic sites that transcription factors bind to when regulating gene expression. Here, complexes of proteins and the DNA they bind are isolated by antibodies, and the target DNA is identified by PCR or sequencing.

After surveying the methods for studying gene expression, let's look at some of their applications.

Cells within a population may exhibit subtle differences in gene expression that can have biological consequences. In this study, researchers placed individual human embryonic stem cells from the same culture into separate wells on a plate. Using quantitative RT-PCR, scientists determined that the expression of *Nanog*—a stem cell "marker"—differed within a sample.

Some researchers investigate whether different isoforms of regulatory proteins function differently. Here, ChIP was applied to human immune cells to identify the binding targets of a protein's "long" and "short" isoforms. Sequencing results showed that some gene targets were only recognized by the short isoform, pointing to potential functional differences.

Finally, reporter assays can be used to evaluate gene regulation mediated by small RNAs, such as microRNAs. As microRNAs can inhibit gene expression by binding to the 3¢ untranslated regions of mRNAs, scientists attached this region from different genes to a luciferase reporter, and introduced each of them into cells along with a microRNA. Gene targets of the microRNA were then identified by looking for cells with decreased luminescence signal.

You've just watched JoVE's introduction to gene expression. We've reviewed major findings in gene expression research, prominent questions and methods in the field, and some current applications. As always, thanks for watching!