Course Name: I think biology

Professor Name: Dr. Sravanti Uppaluri

Department Name: Biology

Institute Name: Azim Premji University

Week:3

Lecture:15

W3L15_Gene Regulation

Welcome to the lecture on gene regulation in the iThink Biology NPTEL course. So in the last class, you would have heard about the central dogma. Central dogma essentially tells us how DNA, which is basically a set of genetic instructions is transcribed into mRNA and this mRNA in turn, encodes for proteins and mRNA is translated into protein. So what gene regulation is, is really thinking about when this process should occur, right?

So what is really gene regulation? So in some cases, there are some genes that always need to be turned on. That means they need to be constitutively expressed. So they should be making RNA and protein all the time. So as long as the organism is alive, there is some need for this mRNA or protein. But in other cases, there are some genes that don't need to be turned on all the time, right, so they need to be regulated. For instance, when there is a change in the environment.

So for example, when we eat food, right, What is our response? We need to produce more enzymes, we need to say produce more insulin so that glucose can be absorbed into our system. So our bodies need to really respond to the environment. And often this response means that some genes need to be turned on or some specific proteins need to be made,right.

Another case where gene regulation is required is during development. What I mean by development is how we go from a single cell to the cells dividing, right. A fertilized egg is a single cell and these cells divide and start to differentiate. That means that you start getting cell-cell differences via a process named differentiation. And this differentiation actually leads to different cell types, right. The different tissue types that we have. So for example, liver cells versus heart cells, right. And these cell-cell differences occur also through gene regulation. So let's have a look first at how genes are regulated in prokaryotes versus eukaryotes. So on the left-hand side is the image of a prokaryotic cell.

It's a schematic, of course, a cartoon image. On the right-hand side is that of a eukaryotic cell. So what you see in the prokaryotic cell is that the cytoplasm is where the DNA is already contained and directly from this DNA within the cytoplasm, mRNA is made and from the mRNA again proteins are made, right.

In the eukaryote, the situation is a little bit more complex, right. So this is within the nucleus now, right. Not within the cell, the cytoplasm. Within the nucleus, the DNA is first made into a pre-mRNA and we'll talk about this in a little bit more detail later on. And this pre-mRNA is processed into a mature mRNA.

This mRNA is transported out into the cytoplasm and then translated into protein. So it turns out that usually in prokaryotes, most genes are switched off. That means that they are repressed. So whereas they are almost always in an active state, whenever the gene needs to be regulated, it is repressed. That means that the mRNA stops being produced. So transcription has to be actively switched off. So the sort of most common form of gene regulation in prokaryotes is repression.

In eukaryotes, it's exactly the opposite. So gene regulation usually occurs through activation. So genes by default are in an off state in eukaryotes and they have to be turned on whenever they're required.

So why does this happen at all? Why is there a difference between prokaryotes and eukaryotes? Well, the answer actually lies in how DNA is packed inside prokaryotes. So in prokaryotes, the DNA is, as I said, directly found in the cytoplasm. So in this cartoon image, you see that actually the DNA is really not, if you look at a zoom in image of it, you see the double helix. It isn't really packed as much as it is in the eukaryotes.

In the eukaryote, you can see very clearly that from the chromosomes, there's a tight packaging, right. They are tightly condensed and they are condensed by being wrapped around these yellow histone proteins, right. And in order for a gene to get expressed right, to be activated, it has first to be unwound from these histone proteins, which makes sense because this is why, you know genes are by default in an off state in eukaryotes, but not so in prokaryotes.

Okay, so just to get an idea of what we mean by gene regulation and how quickly this can happen, suppose you are trying to think about how quickly you can respond to the environment, right. Every organism will have a different rate at which transcription and translation can occur. So how quickly can you make a protein, right? So this table here tells you that for E. coli, right, which is a bacteria, which is prokaryotes, it can transcribe mRNA at a rate of 10 to 100 nucleotides per second, right. So depending on the length of the gene that you need transcribed, it will take a different amount of time to actually get the mRNA molecule.

In turn, if you have, if you look down at the last row in this table, translation of a protein in a mouse, right, goes at the rate of six amino acids per second. So suppose you wanted to, suppose a mouse needed to synthesize insulin, right, you have the mRNA molecule, it has to get translated.

Mouse insulin is about 50 amino acids in length. So you would need 50 into six seconds to produce, so that's about 300 seconds to produce an amino acid protein if you have the processed mRNA molecules, right. So that's about five minutes to produce insulin. So that also plays into how quickly you can respond to different environmental factors, right. So that also means that the cell has to decide which proteins it can produce at sort of on the spot, depending on the reaction or the function of that protein. So generally, you don't need insulin so quickly, so five minutes is a reasonable timeframe for insulin to be produced, right.

So as I said, the other aspect of gene regulation is really when we are thinking about differentiating, that is development, right. So when you have an oocyte meeting a sperm, and this being forming a fertilized egg, as it goes through development, you first have cells that all look very similar, right. Over the course of development, you start to see the human fetus that eventually develops from what is called the blastocyst, right. And the blastocyst initially has this thing called the inner cell mass. So this is a bunch of cells, this is here in the color blue and the nucleus is in orange.

And these cells from the inner mass are eventually, are the precursors that eventually form, say, all of our different tissue systems, right. So the circulatory system from our heart to our, all of our veins and arteries and so on, the nervous system, the immune system, the liver, etc, etc. And these cells are differentiated. So what you see is that they start off from initially the same type of cell and then they start differentiating. So both these examples from the response to the environment, as well as the process of development, both of these are examples where gene regulation is required.

Okay. So the rest of this lecture will primarily focus on this example, where cell differentiation or rather cell regulation is required for cell differentiation. So now let's zoom out a little bit and look at this as a simple bite version. So you have a zygote, right, which is a fertilized egg, and this is a single cell, those cells start to divide, and they form cells that are exactly the same, right. They're copies of each other. But over the course of time, you eventually get cell differentiation. So these cells are no longer similar, right. They start to differentiate, and you eventually get these different types of cells, right. Everything from intestinal cells to stem cells to fibroblast cells.

And what you see from these images is that the cells not only have different functions, but they gain these different functions by having different morphologies, right. So they have really dramatically different shapes and sizes. And they also actually perform dramatically different functions, of course. So how does this happen, right? When you start off with something that's exactly the same, you're making copies that are exactly the same. But over the course of

development, right, through embryogenesis, as the embryo is starting to develop, these cells differentiate. So there are two possibilities, right. We all know that there are genetic instructions.

And in this cartoon, I've shown these genetic instructions as these different shapes, right. So as cells divide, what happens to these genetic instructions? One possibility, right, and this is a possibility that was first proposed by a scientist named Weissman, he said that these instructions, he didn't know about the presence of DNA then, that these instructions could actually be divided as the cells divide. And this division is actually is actually what leads to cells being differentiated, right. And that's what this first model is.

Another idea would be that these determinants, right, or whatever these sets of instructions are, start off in the zygote. And as the zygote divides, right, all of the cells retain all of the genetic instructions. But in each set, a different set of instructions is being followed, right. And it turns out that it is actually the second possibility that is true.

That means that every cell has the same set of instructions, right. And we now know, much after Weissman's time, that the set of instructions are contained within our DNA and our genomes, right, the whole entire set of our DNA. And all of our cells are genomically equivalent. That means that they all contain the same set of instructions. But some instructions are, let's say, preferentially followed in each cell type, right.

Which means that while I have the same set of instructions, a liver cell might decide to produce an enzyme, a particular enzyme that's required, but a cardiomyocyte in your heart may not require the same enzyme, hence would produce something completely different, right. So this is what is referred to as genomic equivalent. So all of our cells have the same set of instructions, right, the same genome, but different parts of that genome are followed in each cell.

So what is that process? And that process is exactly gene regulation, right. So some of these genes are preferentially regulated within each different cell type. How does this happen? And what does it mean for a gene to be regulated? And this is where we come back to our central dogma, right. So DNA, where genes are found, and these genes encode specific proteins, right. How do you go really from DNA to protein? You have to go through the process of transcription and translation. And gene regulation actually occurs everywhere along this process, right, along this process of transcription and translation. And we look at this in a little bit more detail now.

So if you look at the nucleus, right, now again, we are still highlighting the central dogma here, but in a little bit more detail. So we have the DNA. The DNA contains introns and exons. Exons are parts of the DNA that code for protein. Introns don't code for protein, right. Introns also have essential function, and we'll talk about those later on. So now this DNA can be transcribed into RNA. The RNA can then be spliced or cut up and then rejoined where introns are removed.

And this RNA then is exported out into the cytoplasm and translated into protein. Now each level, right, of this process is actually where gene regulation occurs, right. So first of all, not all genes are transcribed. It turns out that this wrapping around the histone proteins that I alluded to earlier, this is an essential part of gene regulation, right. So if DNA is methylated in specific regions, that methylation may or may not allow the DNA to unwrap from the histone proteins. And as we know, this unwrapping is really essential for RNA polymerase to come and bind to the protein, to the DNA, which then allows transcription to occur. So differential gene transcription is one sort of gate through which a gene has to pass through before it can be expressed.

Now once it's transcribed, it has to be processed. Again, a cell may choose to keep this RNA nascent and not have it, not process it at all, so not splice it immediately, but wait for it to be spliced at the right time in the right place during development. Once it's spliced, the cell has to decide whether or not it wants to export it out into the cytoplasm. And once it's exported, the cell has to decide whether it can, it should be translated or not. So this is what we are referring to when we talk about selective translation.

It means that a gene may or may not be considered expressed until it has actually been translated, right. Once translation occurs and the protein is formed, the protein is still inactive until some modifications have been made to that protein. So again, this is another gate through which genes have to cross through in order for them to actually be considered expressed, right, for the protein, the final product of gene expression to occur. So this is really the essence. All of these steps are what have to be followed and these gates really have to be crossed before a gene can really be expressed. And it's at all of these levels that gene expression is controlled at, you know, different combinations and in different cell types for us to actually get the tissue types that we see in complex multicellular organism.

Now I just want to give you a picture. We've really, in the previous slide, we gave you a very coarse-grained view of it. I just want to give you a picture of one process which is for transcription, right. So it isn't just that transcription is controlled by DNA methylation and histone unwrapping. Transcription can also be controlled through different kinds of proteins. So even when the DNA is unwrapped, a transcription factor first has to bind and enhance our sequence that is close to the gene that we want to express.

So here this gene is highlighted in yellow. Once a transcription factor binds to the enhancer sequence, this has to now recruit an RNA polymerase. Now remember both the transcription factor and RNA polymerase are proteins themselves, right. So the expression of these proteins also is regulated. So you can see that it's quite a complicated system where everything feeds into the next process, right.

So the production of the transcription factor itself has to be regulated. And once it's produced, the binding to the enhancer sequence has to be regulated. Often it isn't just one transcription factor binding to an enhancer sequence. It's multiple transcription factors binding to multiple enhancer sequences. And depending on the combination of enhancer sequences that bind, it might promote the expression of a different gene.

So I start to show you that this system is really quite complex and fascinating. So once a transcription factor binds, let's say now in this cartoon that we just have one transcription factor binding, this then allows the recruitment of RNA polymerase to come and bind to the promoter site. And as I said earlier, the promoter site is where RNA polymerase binds. And this now starts to help unravel or separate the double strands of the DNA, which then can lead to transcription of the gene. And once we've transcribed the gene, again, the same process must begin all over again, where you have mRNA molecule, the mRNA molecule has to be spliced and processed.

Again, one more gate, this transport to the cytoplasm, another gate for gene regulation, and then protein translation and modification. All of these are gates that the gene has to go through in order for it to be really expressed,right. So this actually brings us back to the central dogma, and what I want to highlight, and I hope you can see from what we've just discussed in this lecture is that the central dogma probably needs a little bit of addition in the sense that it isn't really a linear process where you get DNA that is the set of instructions that is transcribed to RNA, which then makes protein.

Rather, all of these feed into each other,right. So the protein that, proteins also regulate DNA. And that was the example of the transcription factor that I gave you,right. RNA can also be transcribed into non-coding RNAs. So those RNAs never actually make protein, but they can regulate the production of protein. Finally, proteins through RNA splicing, right, can be produced in different isoforms.

So it isn't just that RNA, every RNA just makes one protein. RNAs can make different, the same RNA molecule before it is spliced can actually be a precursor for a whole family of proteins. And again, these proteins can regulate translation. They can also regulate DNA transcription. And so you can see that there is a lot of feedback in the system and the complexity of it is really quite mind boggling.

And if you put this in the context of evolution, which we will study a little bit later on in this course, I think our fascination with biology will really only increase manyfold. In the next lecture, you will actually be looking at non-coding RNAs and their different roles in the cell.