Video Tutorial 14.1: Eukaryotic transcription networks

The sequence of a genome provides the information that directs cells to particular fates and functions, but it does not provide the strategy that underlies how this information is used. These strategies for humans and for several model organisms are being unraveled by a consortium consisting of dozens of laboratories, hundreds of investigators, and thousands of individual experiments that together make up the ENCODE projects—the Encyclopedia of DNA Elements.

Let's look at some of the transcriptional regulation in Drosophila melanogaster as reported in the Fly ENCODE project. We go to modENCODE.org, and then to View Fly Regulatory Network. This takes us to this diagram.

Each of the green circles represents a Drosophila transcription factor—76 different transcription factors are shown. These 76 transcription factors represent about 2 or 3% of the total transcription factors encoded in the fly genome, but the patterns of interactions that they illustrate likely are general. The red circles are microRNA genes; for now, I am going to turn off the interactions between transcription factors and microRNAs, although that is something to explore yourself.

For each transcription factor, its targets were found by chromatin immunoprecipitation, as discussed in the chapter. The interaction between the transcription factor protein and its targets are illustrated by green arrows, pointing from the transcription factor to its targets. Think of the arrows as giving and receiving instructions. Note that only the target genes that encode other transcription factors are being shown. Nearly all of these transcription factors regulate other genes—genes that encode enzymes, signaling molecules, structural proteins, and so on—but this diagram is only looking at the strategy by which transcription factors regulate the expression of each other.

The interactions were plotted computationally to show the regulatory interactions, and this hierarchy emerges. This particular plot shows five tiers in the hierarchy, but I don't think that the number of tiers or even the strict assignment of a particular transcription factor to a tier is all that meaningful; the important point is the overall picture that arises.

Let's look at the transcription factor known as hairy—shown by the h in the top tier.   When we click on it, we see all of the interactions hairy has with other transcription factors, which illustrates that hairy is clearly a central player in the organization of Drosophila transcription. All of the genes that directly interact with hairy remain highlighted while the ones that do not directly interact with either are dimmed out.   Note how many arrows emerge from hairy and point to other genes, about 50 of them—these are genes encoding other transcription factors whose expression is directly regulated by hairy. These ended up either in the same tier as hairy or below it in the hierarchy.

Note also how many other transcription factors regulate the expression of hairy. The picture is a bit crowded but there are about a dozen genes encoding transcription factors that directly regulate the expression of hairy. One of these is hairy itself—in other words, hairy is an example of auto-regulation, of which there are many in this diagram.

Nearly all of the genes that regulate hairy are in its same tier or in the tier immediately below it. Many of these genes not only regulate the expression of hairy but also are in turn regulated by hairy. The Kruppel gene, abbreviated by Kr, is an example. Kruppel is a gene important in Drosophila embryogenesis. Notice that it is regulated by hairy but also that it regulates hairy.

The hairy protein binds to the regulatory module of Kr and regulates its transcription. The Kr protein binds to the regulatory module of hairy and regulates its transcription. This type of cross-regulation between genes and gene products is what gives rise to the particular pattern of transcription factors that governs Drosophila development.

The diagram shows that the interactions occur but it does not provide information on the dynamics of the interactions. We can't tell from this diagram if hairy is expressed before Kruppel, or if Kruppel is expressed before hairy, or if they are expressed at the same time. But that can be sorted out using the links to the papers that describe the interactions with hairy, and it serves to direct more specific questions.

You can click through any of these genes and see examples of the hierarchy at work. Instructions come primarily from genes in the same tier or in the tiers above. Instructions are passed down to the genes in the same tier or in the tiers below. This is not a strict rule but it certainly covers most of the hundreds of interactions that are depicted here.

Let's look at one more example. At the lowest tier is a transcription factor known as en or engrailed. If we click on en, we see that it regulates two other transcription factors—itself and the inv gene in the tier immediately above it. The relationship between en and inv is worth looking into but can be left to you to look up in you are interested in thinking about gene families a bit more. Most of the targets of en turn out to be the effector genes rather than other transcription factors. But en is the target of regulation by more than 30 other transcription factors.

We have not included the microRNAs in this picture, but they are their own story. In fact, each of these interactions is its own story, and they suggest many more possible experiments to think about. With only a little imagination, we can envision hundreds of PhD thesis projects and publications embedded in this diagram.

An important goal for genetic analysis of an organism is to have the sequence of its genome. But the ENCODE projects show clearly that the sequence of the genome is an intermediate goal when it comes to thinking about biology.