Class 18: Transcriptomics

Reading assignment

- Dale From Genes to Genomes: Chapt 11.

Classroom activity (limit 45 minutes)

1. Dale *Genes to Genomes* Chapter 11 describes several methods for analyzing the transcription of large numbers of genes. On a separate sheet, make a list of the techniques indicating for each (1) how quantitative the method is, (2) how large a fraction of the genome it can analyze, (3) how the measured transcription activities are assigned to the gene being transcribed and (4) major technical advantages or disadvantages of each technique.

   Methods: differential screening, subtractive hybridization, various microarrays and RNA seq (and some others). The first two are qualitative, identify DNA clones based on how much they're transcribed and focus on specific genes rather than the whole transcriptome. Microarrays provide data across the transcriptome but only find genes already identified and the results are often at best qualitative. RNAseq allows very specific, local data on RNA abundance so it gives a much higher resolution result and can be quite quantitative. The first two are easy to perform but provide limited data and miss anything but very abundant RNAs. the latter two provide broader data and RNAseq is clearly better at finding all RNAs, even low abundance, with sufficient depth of data. It is taking over the experimental field.

2. You are a professor at a small Midwestern college and are planning to design a laboratory course for undergraduate students that will be discovery based rather than cookbook; you want to design a course where students can explore how much variation there is in the way gene transcription responds to changes in the cell's environment. Your Chair says that the budget for this class is rather limited so you need to conserve your money. Describe which of the techniques you would be most likely to use in the laboratory and explain why you would make that choice.

   A small midwestern college cannot afford RNAseq but might be able to afford some kinds of microarrays. An array that can be purchased from a supplier would be the best alternative but still is expensive. It would be much more cost effective to focus on some of the earlier generation strategies like subtractive hybridization and differential screening. Those are quite inexpensive but still provide information on mRNA abundance. Being exhaustive isn't so important for this situation.

3. You've just been hired to teach a similar course at a large Eastern private research university and you have the same idea. Your Chair at this institution says that your budget, though limited, is generous. You want to impress the Chair by creating a really impressive course but you are also want to save a few dollars. Your hope is to have students look at large scale differences in transcription across the genome in response to various treatments that they will have to determine. What technique would you chose in that case? Explain why.

   The large, Eastern private university may have lots of money and so it may be no object. They could easily use some version of RNAseq to do the analysis. It might make sense to use an organism with a smaller transcriptome, like yeast, rather than use mammalian systems. The second approach of using microarrays is also attractive because, while expensive, it is easily scaled to a large classroom and students would get a real-life experience.

4. You are applying for funding from the National Institutes of Health for funds to do the first genome-wide studies of transcription in a group of closely related bacterial species but the genome sequences of these species' genomes have not been completed. What strategy would you adopt for this study? Explain your choices.

   Probably the best idea would be to use second-generation sequencing to complete the genome sequence of several of these bacterial species and then adopt microarray or more probably RNAseq approaches to study their transcription. Using an earlier technology would probably be a non-starter because the results would be far less than transcriptome-wide and would not be quantitative.

5. The chapter also considered genome-level analysis of protein levels, which is important since many genes are regulated at the translational level. Compare the advantages and disadvantages of two-dimensional electrophoresis and mass spectrometry.

   Two-dimensional electrophoresis provides a quick, qualitative look at protein levels and can allow researchers to identify different forms of each protein. It is far less quantitative than mass spectrometry and cannot define the differences between the multiple forms of a protein that might, for example, be created by post-translational modification. Mass spectrometry is arguably even less able to provide quantitative results unless special techniques are used to ease comparison between samples but is very accurate about the structure of the proteins, especially their post-translational modification. Mass spectrometry is also much more expensive than two-dimensional electrophoresis by several orders of magnitude.

6. The chapter introduces two methods of analyzing protein-protein interactions: yeast two-hybrid analysis and phage display. What are the advantages and disadvantages of each procedure? Is it possible for either to identify interactions between proteins that do not normally interact in vivo? Why or why not?

   Yeast two-hybrid analysis tests the interaction by fusing a promoter binding protein to one of the proteins to be tested and a transcription activation domain to the other and measuring stimulation of transcription. This allows testing the interaction in vivo, preserving any protein modifications and giving a quantitative answer (stronger interactions give higher activity). It also allows selection of rare interactions from a large library. The disadvantage is that the interaction must occur in the nucleus and the proteins might have strong cytoplasmic localization regions. Also, the proteins might interact as part of a large complex, which would not form at the promoter. Phage display is an in vitro method that fuses one protein to the M13 bacteriophage coat and requires an interaction with a second protein immobilized on the surface of a test tube. A large number of phage fusions are tested and any that does interact is purified and forms a plaque when subsequently infected in bacteria. The advantage is that a very large number of interactions can be tested allowing selection of a successful protein-protein interaction. The disadvantage for a eukaryotic protein is that it might not be properly post-translationally modified blocking interaction.