Assignment 1 – Experimental Approaches

The first task for this work would be the integration of a reporter gene into a mammalian genome. I would choose an inducible promoter, which will be useful in probing the effects of transcriptional activators on expression bursts. For the reporter gene itself, I would choose a gene that is not native to the host cell so as not to interfere with normal cellular function.

A retroviral infection could integrate any gene into a random chromosomal location. I would then check for stable integration by looking for the insertion (by PCR of purified DNA) over several generations. One concern is that the integration could disrupt the cell, especially in the function of genes involved in regulation. Sequencing the region around the insert (by using outward facing primers that bind in the insert) would determine exactly where in the chromosome the reporter gene resides. We could then determine if the nearby and likely affected genes are likely to disrupt normal cellular functions.

The primary experiment of this paper is a measurement of mRNA abundance in single cells. Standard bulk, population assays for RNA (such as quantitative RT-PCR) would not work. In order to count individual mRNA molecules, I would use RNA FISH (fluorescence *in situ* hybridization).

A large set of singly fluorescent labeled oligonucleotide probes for an individual gene are introduced into a fixed cell, and only probes that aggregate to a high enough degree via cooperative binding to a complementary mRNA will be visible by fluorescent microscopy. Thus one could count fluorescent spots that correspond to individual mRNA molecules in a cell at a single time point.

This protocol requires samples to be fixed, so mRNA counts can only be taken as a snapshot. No synthesis rates can be determined, and inferences on expression burst timing must be made from comparing mRNA counts across cells. Assembling the mRNA counts from many individual cells would allow us to determine a distribution that could tell us about the randomness of expression.

I would create a set of probes for the reporter gene, one control natural gene near the reporter (close and on the same chromosome), and one control natural gene on a separate chromosome. Comparing mRNA counts for these three genes would allow us to determine if expression is related to genomic locus becoming more or less accessible for transcription. If there is a correlation between the expression of the close genes (in individual cells) but not the genes which are on separate chromosomes, we have evidence for upregulation of a genomic locus.

To control for the effect of genes that are functionally regulated together, we could compare the correlation of high and low counts in single cells with the distributions of each mRNA across many cells. If the population (across many cells) average mRNA count for the two close genes is very close, then individual cell observations of similar counts do not tell us anything about locus activity.

Observing mRNA counts for the natural genes alone would tell us the same information that we could acquire from a constitutively active reporter, but modulating reporter expression with an inducible promoter would tell us how burst statistics are affected by transcriptional activators. The same FISH assay could be conducted on cell populations exposed to various levels of activators.

To assess the effect of protein degradation time on the ultimate protein expression levels, we could compare protein counts of the unstable, normal, and extra stable versions of the reporter. GFP, for example, has been engineered to have a spectrum of degradation times. Using single cell GFP fluorescence as an indicator for the count of GFP proteins in a cell, comparison of the variation of GFP counts across cells with the variation of *gfp* mRNA counts would tell us if variation is buffered at the protein level. Repeating this experiment for stable or unstable variants of GFP would then tell us the degree that variation (from burst effects) are attenuated by the stability of the protein.

The stochastic model of transcription activation could be based on a simple set of reactions: (1.)gene transfers from inactive to active state and (2.)back again, and (3.) active gene transcribes mRNA at a rate proportional to level of transcription activators. I would determine parameter values by trying to match simulated expression data to experimental counts.

7.342 Systems and Synthetic Biology: How the Cell Solves Problems
Fall 2010