Connecting mechanistic models of transcriptional regulation and single-cell gene expression data

Alvaro Sanchez[1], Jané Kondev[2]

We present an analytical theory of stochastic gene expression, which is able to predict mean expression and noise patterns for promoters of arbitrarily complex regulatory mechanism [1]. Our theory makes it possible to quantitatively connect regulatory mechanisms with single cell gene expression data. The theory has been applied to the GAL1* promoter in yeast, showing excellent agreement with experimental data and simulations [1,2]. The theory offers insights into potential mechanisms to independently control the mean gene expression and the cell to cell variability. Such independent control of noise and mean expression is predicted to occur in the GAL1* promoter in yeast.

I. Purpose

UNLIKE traditional gene expression experiments, single cell gene expression experiments provide not only the mean, but also the whole distribution of gene expression in a population of cells (i.e. the cell to cell variability or "noise" in gene expression). These kind of experiments call for the development of analytical tools that help to connect proposed regulatory mechanisms with experimental data, test hypothesis, and guide the formulation of new experiments. So far, evaluation of these regulatory mechanisms and its comparison with the single cell data, has been achieved mostly through stochastic computer simulations, except for those cases in which the regulatory mechanism was simple enough to be reduced to a two-state (active/inactive) mechanism, in which transcripts are assumed to be made at different but constant rates at the active and inactive states. An analytical derivation of the moments of the steady state probability distribution of mRNA produced by this mechanism exists [3,4]

We have recently developed a stochastic analytical theory that generalizes this to promoters with any number of regulatory states (depending both on which transcription factors are bound where, and the general state of the promoter with respect to the conformation of DNA), calculating all the moments of the steady state probability distribution. The theory was shown to agree very well with in vivo gene expression data and stochastic simulations, when we applied it to a specific example, the GAL1* promoter in yeast [2]. Furthermore, the theory suggests a potential mechanism to independently tune the mean expression and the cell to cell variability in complex promoters regulated by more than one transcription factor. The independent control of noise and mean is predicted to occur in the GAL1* promoter.

Using our theory, we investigate the noise characteristics of a few typical cis-regulatory motifs, whose mean response to transcription factor concentration is well characterized [5]. We explore the role of cooperative binding of transcription factors, number and strength of transcription factor binding sequences on the DNA, and regulation by DNA looping in the generation of noise in gene expression.

II. Conclusions

Our theory incorporates stochastic effects and is capable to connect mechanistic models of transcriptional regulation with single-cell gene expression data. The theory makes a number of experimentally testable predictions on the behavior of the cell to cell variability as a function of transcription factor or inducer molecule concentrations for specific promoter systems. We expect this theory will be used to test proposed mechanisms of gene regulation and guide new experiments.

References

[1] Sanchez A, Kondev J (2008) "Transcriptional control of noise in gene expression". *Proc Natl Acad Sci USA* **105,** 5081-5086.

[2] Blake WJ et al (2003) "Noise in Eukaryotic gene expression" *Nature*. **422**:633–637

[3] Kepler TB, Elston TC (2001) "Stochasticity in transcriptional regulation: Origins, consequences and mathematical representation" *Biophys J*. **81:**3116-3136

[4] Paulsson J (2005) "Summing up the noise in gene expression". *Nature* **427**: 415-418

[5] Bintu et al (2005) "Transcriptional regulation by the numbers: Models" *Curr Opin Genet Dev* **15:**116–124

Acknowledgements: This work was funded by NIH grant XX00000. This footnote is optional.

[1]Graduate Program in Biophysics and Structural Biology, Brandeis University, Waltham MA. E-mail: aasanchez@brandeis.edu

[2]Department of Physics, Brandeis University, Waltham, MA. E-mail: kondev@brandeis.edu