Interferon-alpha induced transcriptional bursting in space and time

C.W. Seitz

June 12, 2022

1 Introduction

Interferon signaling plays a complex role in the innate immune response following pathogen detection. Interferon molecules bind to cell surface receptors and initiate a signaling cascade leading to the transcriptional regulation of hundreds of IFN-regulated genes (IRGs). Recently, a new wave of interest in the precise mechanisms controlling transcription of IRGs has developed due to the advent of cytokine-based cancer treatments and implication of inteferons in autoimmune diseases such as type-1 diabetes. Genome-scale sequencing assays have revealed that interferon-alpha can induce chromatin remodeling as well as changes in transcriptional and translational regulation. In particular, ATAC-seq of human beta cells following interferon-alpha treatment has shown that over 4400 gained open chromatin regions (OCRs) formed two hours follwing cytokine treatment. These OCRs have been defined as putative distal regulatory elements for ISGs - a definition which is evident in the transcriptional up-regulation of genes within 40kb of an interferon-induced OCR. These observed changes in chromatin accessibility and upregulation of ISGs raise important questions regarding the spatial and temporal properties of gene expression. The development of OCRs following interferon stimulation may facilitate transcription; however, the spatial and temporal dynamics of ISG expression remain unclear. A rigorous characterization of the timescale and mode of transcription (e.g., non-constitutive or constitutive) would contribute to our understanding of the spatiotemporal kinetics of the transcriptional response to interferon stimulation.

2 ISG transcript-count distributions are inconsistent with Poissonian gene expression

Intracellular biological reactions are inherently stochastic which manifests as cell-to-cell variability, even in isogenic populations of cells. Furthermore, the structure of this population noise provides important insights into the underlying biochemical processes. The classical view of constitutive gene expression in which gene products accumulate in a Poisson-like process has been challenged

by recent observations of episodic transcription. Thus far, this so-called transcriptional bursting has been attributed to chromatin remodeling as well as modulation of transcription by upstream transcription factors []. In the common transcriptional bursting model, a promoter toggles between a transcriptionally inactive state and an active state. This switching behavior generates super-Poissonian variation in cellular RNA counts, as opposed to Poissonian regimes in which transcription occurs at a constant rate. In several cases, a two-state model does not fit the data accurately and multi-state models are required. Transcriptional paradigms with many states may provide a richer set of responses to stimulation.

Deviations from constitutive expression are evident in the Fano factor.

3   Testing the ergodic hypothesis

Discuss ergodic and non-ergodic regimes of transcriptional bursting. This determines the statistical equivalence of ensemble snapshots and time-series data

4   Diffusion of nascent RNA in live-cell nuclei

Most models thus far have focused on inferring kinetic parameters at a genomic locus with the fate of nascent transcripts left unaddressed. This could be, in part, because evidence has surfaced that transcripts in the nucleus primarily exhibit Brownian motion on their way to the nuclear pores at the nuclear membrane (Singer). At the same time, the pulsation of RNA production followed by Brownian motion makes the spatial arrangement of RNA near the transcription site non-trivial, due to the overlap of multiple distributions, complexities of RNA elongation and cleavage, etc.. This also is in conflict with the nuclear to cytoplasmic export model used in compartment models of gene expression (Munsky). A first goal is to verify that motion in the nucleus is indeed Brownian while transcription is bursty in live cells (Quintero). These can be analyzed simultaneously, determining the validity of model assumptions along the way. Quintero et al. have published the dynamics at the transcription site for 14 cells imaged at 7Hz, which could be quite useful for testing ergodicity of transcriptional bursting. Two of these cells show 50 nascent transcripts which can be tracked and their dynamics characterized, providing insight for spatio temporal models. Their data also correlates RNAPII binding with RNA elongation, which may be helpful information for dynamical models.

If RNA copy number is an ergodic process, is the diffusion also ergodic? What type of diffusion do we observe and why?

5   Multi-state transcriptional models

This reminds me of the idea of 'transcriptional logic'. If the transcriptional state of the cell is non-binary, there might be a more robust set of responses.

Evolution could have shaped a deeper alphabet for promoter activation. Can we determine how many states there are from the live cell data?

6 Bayesian inference for spatial models of gene expression

In-vivo data is ideal for kinetic parameter inference as it requires minimal assumptions and is compatible with mature dynamical modeling techniques, such as hidden Markov models (Rattray). However, studying transcription in-vivo requires a much more intensive biological component with respect to in-situ methods. Also, in-situ methods are scalable in the sense that many unique RNA species can be fluorescently tagged simultaneously, providing a substrate for the study of coregulation of genes.

Due to these complications, kinetic inference using in-situ datasets has been well-studied in the literature, particularly by fitting models of promoter switching to ensemble snapshot time sequences (Munsky). Bayesian inference of kinetic parameters has become quite popular for in-situ datasets as this class of methods does not require detailed time-series data of transcription. Under the so-called ergodic assumption, the statistics of ensemble snapshots and time-series data are identical and useful for parameter inference. One can imagine that, when cells are indistinguishable in this way, the probability density over transcript counts is the same for the ensemble as for the time-resolved cell. In any case, computing the likelihood function for these data is generally intractable, and we have to make use of supplementary methods. Munsky et al. have performed a nice comparison study for fitting models where the likelihood which has been reduced to first and second order moments or computed explicitly by solving a chemical master equation (CME). Samples from the posterior on kinetic parameters are found in all cases via a MCMC algorithm, giving parameter uncertainties. Likelihood-free methods such as approximate Bayesian computation (ABC) avoid computing the likelihood, instead defining a notion of distance between the observed data and simulated trajectories using the Gillespie algorithm.

Compartment models define promoter states and export of RNA out of the nucleus. Inference on compartment models is performed by using the finite state projection of the chemical master equation to compute the likelihood of parameters. RNA export rates are not diffusion-based in existing models, can we do this?

Three items are cited: *The LATEX Companion* book Goossens et al. [1993], the Einstein journal paper Einstein [1905], and the Donald Knuth's website Knuth. The LATEX related items are Goossens et al. [1993], Knuth.

References

Michel Goossens, Frank Mittelbach, and Alexander Samarin. *The LATEX Companion*. Addison-Wesley, Reading, Massachusetts, 1993.

Albert Einstein. Zur Elektrodynamik bewegter Körper. (German) [On the electrodynamics of moving bodies]. *Annalen der Physik*, 322(10):891–921, 1905. doi: http://dx.doi.org/10.1002/andp.19053221004.

Donald Knuth. Knuth: Computers and typesetting. URL `http://www-cs-faculty.stanford.edu/~uno/abcde.html`.

A   Kolmogorov's forward equation

B   Photon statistics for fluorescence microscopy

C   Single molecule localization: The Cramer-Rao bound

Discuss parameter uncertainty, Fisher information and CR bound for RNA localization. This is important for achieving accuracy in particle detection and diffusion models