Cell

Real-Time Kinetics of Gene Activity in Individual Bacteria

Ido Golding,[1,*] Johan Paulsson,[2,3] Scott M. Zawilski,[1] and Edward C. Cox[1,*]
[1] Department of Molecular Biology, Princeton University, Princeton, NJ 08544, USA
[2] Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, United Kingdom
[3] Present address: Department of Systems Biology, Harvard University, Boston, MA 02108, USA.
*Contact: igolding@princeton.edu (I.G.); ecox@princeton.edu (E.C.C.)
DOI 10.1016/j.cell.2005.09.031

SUMMARY

Protein levels have been shown to vary substantially between individual cells in clonal populations. In prokaryotes, the contribution to such fluctuations from the inherent randomness of gene expression has largely been attributed to having just a few transcripts of the corresponding mRNAs. By contrast, eukaryotic studies tend to emphasize chromatin remodeling and burst-like transcription. Here, we study single-cell transcription in *Escherichia coli* by measuring mRNA levels in individual living cells. The results directly demonstrate transcriptional bursting, similar to that indirectly inferred for eukaryotes. We also measure mRNA partitioning at cell division and correlate mRNA and protein levels in single cells. Partitioning is approximately binomial, and mRNA-protein correlations are weaker earlier in the cell cycle, where cell division has recently randomized the relative concentrations. Our methods further extend protein-based approaches by counting the integer-valued number of transcript with single-molecule resolution. This greatly facilitates kinetic interpretations in terms of the integer-valued random processes that produce the fluctuations.

INTRODUCTION

Gene expression involves a succession of probabilistic events: DNA continually undergoes conformational changes, repressors and transcription factors randomly bind and fall off their operators and promoters, and transcription and translation are complex at the levels of initiation, elongation, and termination (Kaern et al., 2005). Even in a hypothetically constant and homogeneous intracellular environment, this complexity would produce random fluctuations in the number of mRNAs and proteins per cell, constituting "noise" that cells must either exploit, learn to live with, or overcome using various noise-suppression mechanisms.

The last three decades have seen numerous probabilistic models of gene expression. Most fall into one of two categories. Some focus on how spontaneous small-number Poisson fluctuations in mRNA levels enslave the levels of their encoded proteins, possibly through bursts of translation (Berg, 1978; McAdams and Arkin, 1997; Rigney, 1979a, 1979b; Swain et al., 2002; Thattai and van Oudenaarden, 2001). Others instead focus on how mRNA fluctuations in turn are enslaved by random changes in gene activity and possible bursts of transcription (Blake et al., 2003; Kepler and Elston, 2001; Peccoud and Ycart, 1995; Raser and O'Shea, 2004; Sasai and Wolynes, 2003; Tapaswi et al., 1987).

The corresponding experimental interpretations have been similarly divided between these two categories. The first quantitative study, using a single GFP reporter in *Bacillus subtilis*, interpreted the results in terms of small-number mRNA fluctuations and translation bursts (Ozbudak et al., 2002). A second *E. coli* study used correlations between dual fluorescent reporters and similarly interpreted the inherent randomness of gene expression (termed "intrinsic noise") in terms of small-number mRNA fluctuations (Elowitz et al., 2002; Swain et al., 2002). In eukaryotes, on the other hand, the first single-reporter study in *Saccharomyces cerevisiae* suggested a substantial contribution from chromatin remodeling, producing quantal transcription bursts (Blake et al., 2003). A follow-up dual-reporter study (Raser and O'Shea, 2004) in *S. cerevisiae* greatly elaborated on these results and also suggested a substantial contribution from transcriptional bursting. Because chromatin remodeling is eukaryote specific, this has been suggested as a possible difference between these two domains of life (Blake et al., 2003).

A difficulty when analyzing the randomness of gene expression is that existing single-cell techniques only allow accurate quantitation of protein levels, while mRNA fluctuations are at best estimated qualitatively (Le et al., 2005; Tolker-Nielsen et al., 1998). Another difficulty is that single molecules of GFP are generally undetectable in vivo due to background fluorescence. With rare exceptions (Rosenfeld et al., 2005), fluorescence data therefore do not report the actual

Every cell has different activity in gene expression

Questions and Background

Central dogma

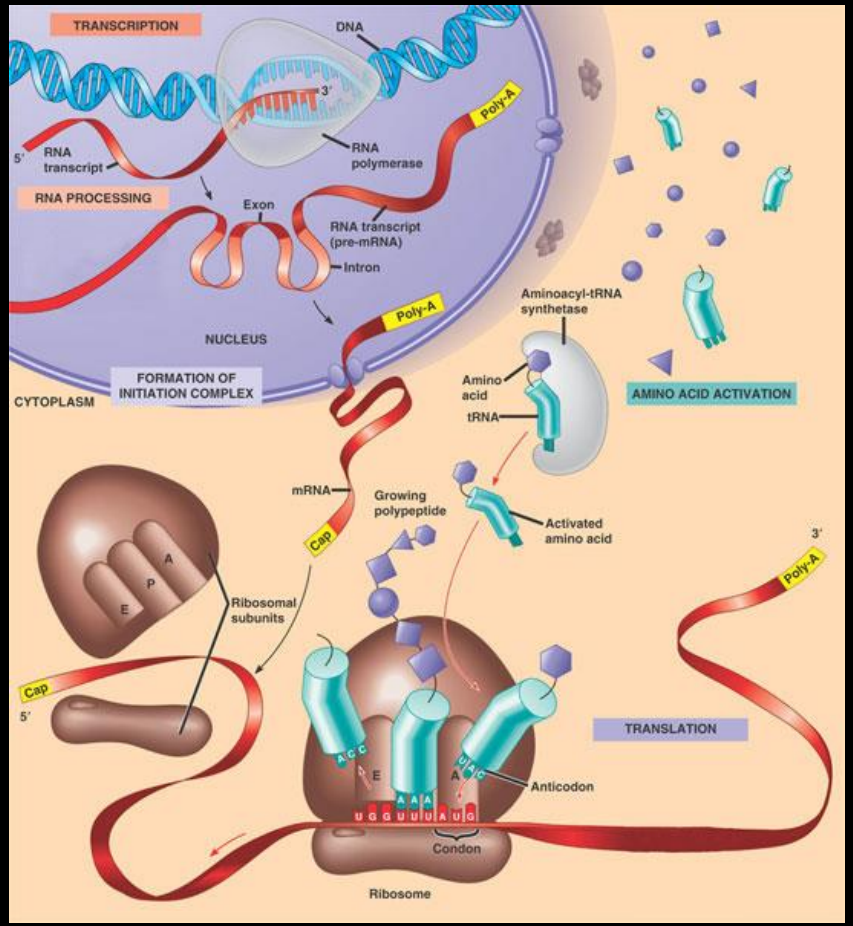
DNA

↓ transcription

RNA

↓ translation

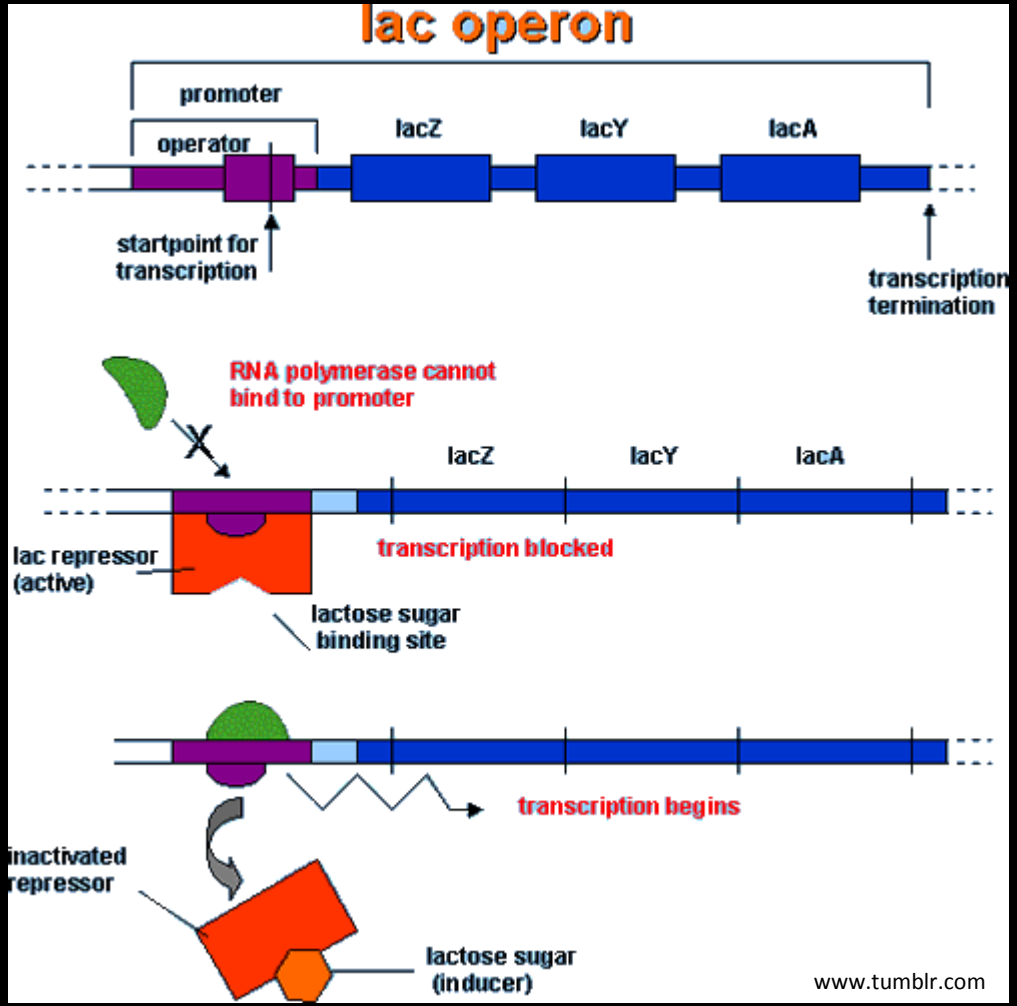
Protein



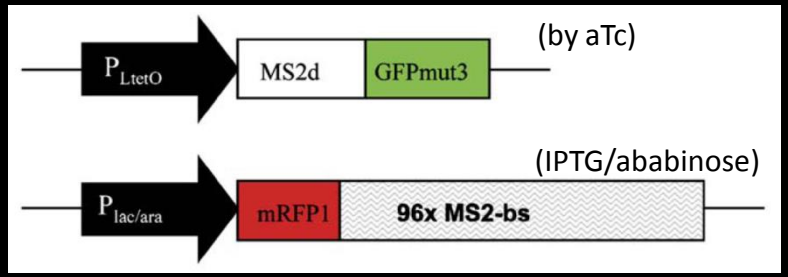
http://compbio.pbworks.com/f/central_dogma.jpg

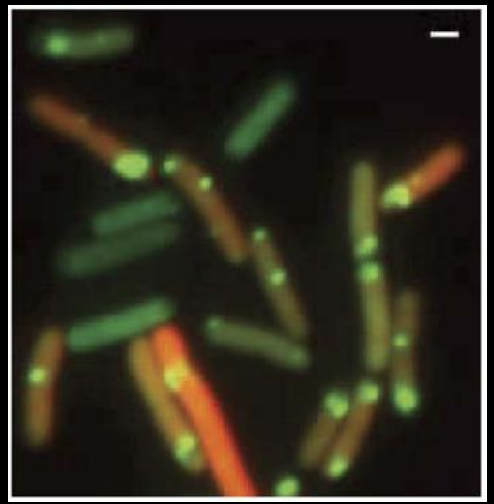
Questions and Background

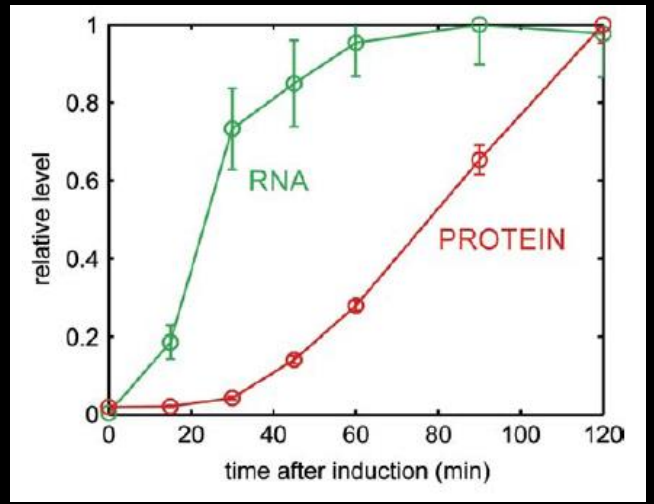
transcription and gene regulation

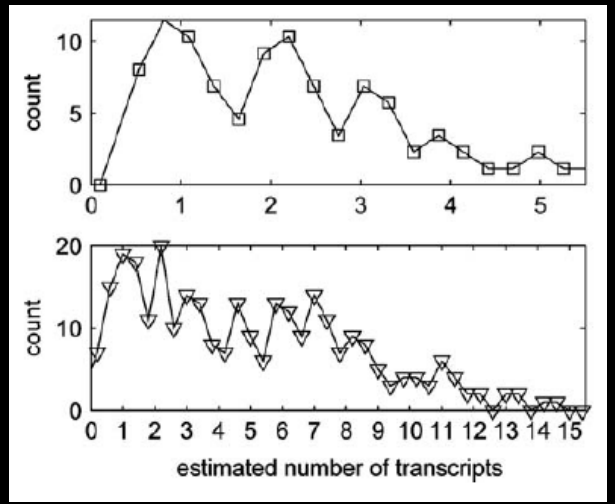


The method


(by aTc)
P_{LtetO}
MS2d
GFPmut3
(IPTG/ababinose)
$P_{lac/ara}$
mRFP1
96x MS2-bs




relative level
time after induction (min)
RNA
PROTEIN
0
0.2
0.4
0.6
0.8
1
0
20
40
60
80
100
120

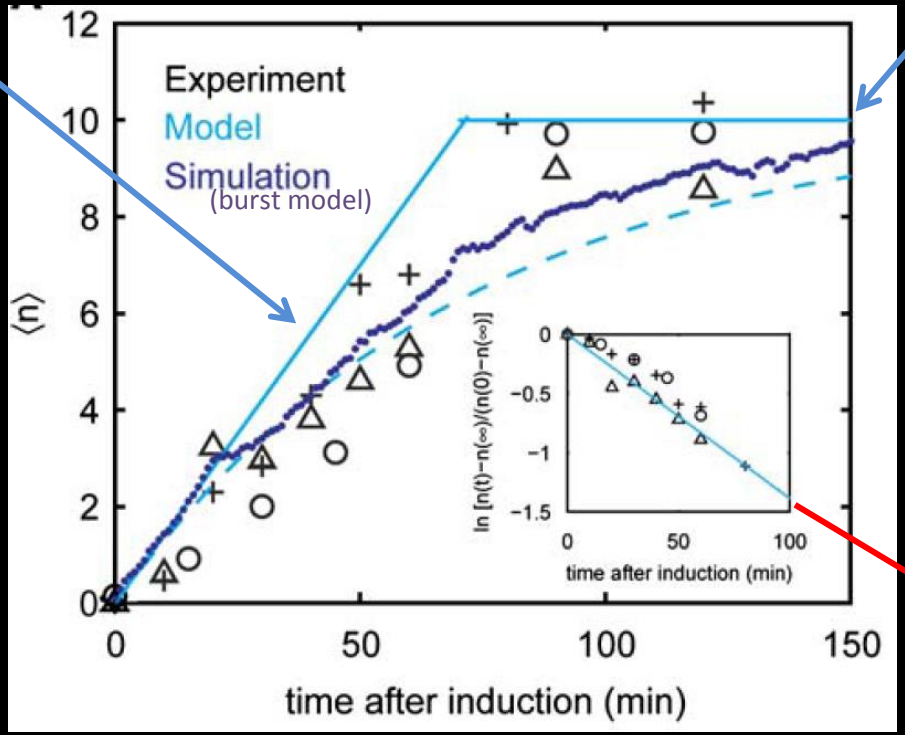

count
estimated number of transcripts
0
5
10
20
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

The method

$$n\langle t\rangle = k_1 t$$

$$k_1 = 0.14\ \text{min}^{-1}$$

+, Δ, ○: three independent data sets



$$n\langle t\rangle = \frac{k_1}{k_2}$$

$$\frac{d}{dt}\langle n\rangle = k_1 - k_2\langle n\rangle$$

$$\langle n(t)\rangle = \frac{k_1}{k_2}\left(1 - e^{-k_2 t}\right)$$

$$\frac{\langle n(\infty)\rangle - \langle n(t)\rangle}{\langle n(\infty)\rangle - \langle n(0)\rangle} = e^{-k_2 t}$$

$$k_2 = 0.014\ \text{min}^{-1} \sim (71.4\ \text{min})$$

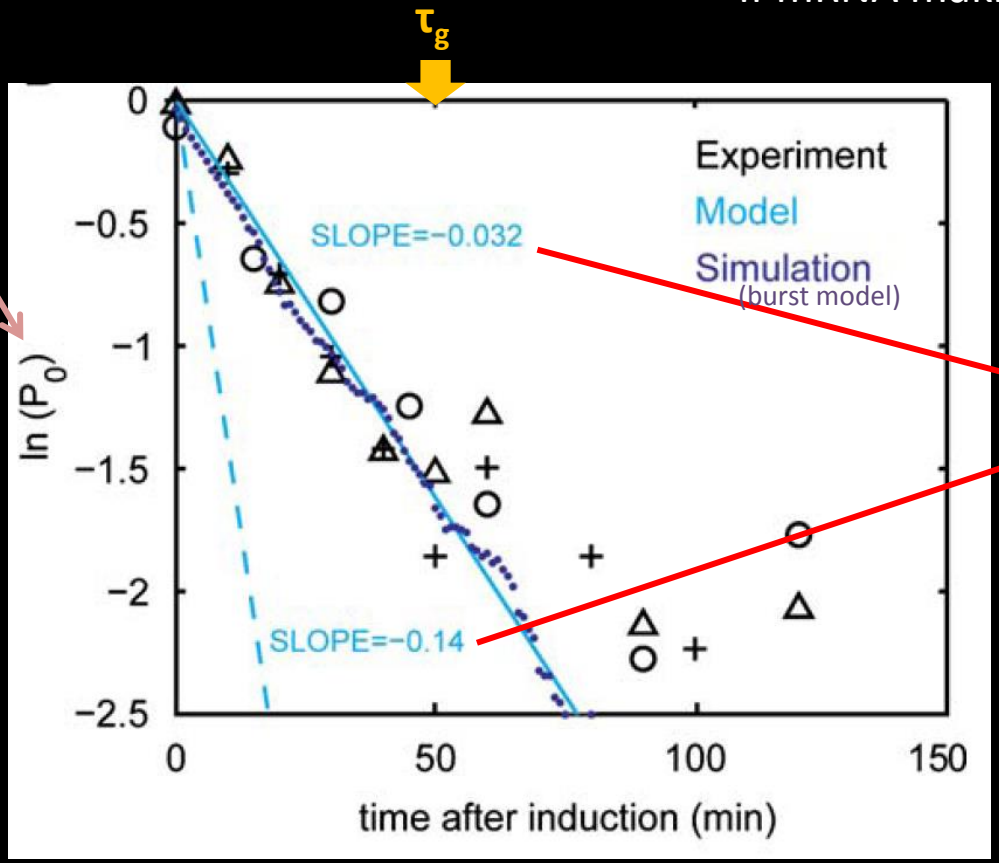
$$\tau_g \sim 50\ \text{min (the cell generation time)}$$

Testing the Poisson hypothesis 1

If mRNA making is the Poisson process,

$$P_0(t) = e^{-k_1 t}$$

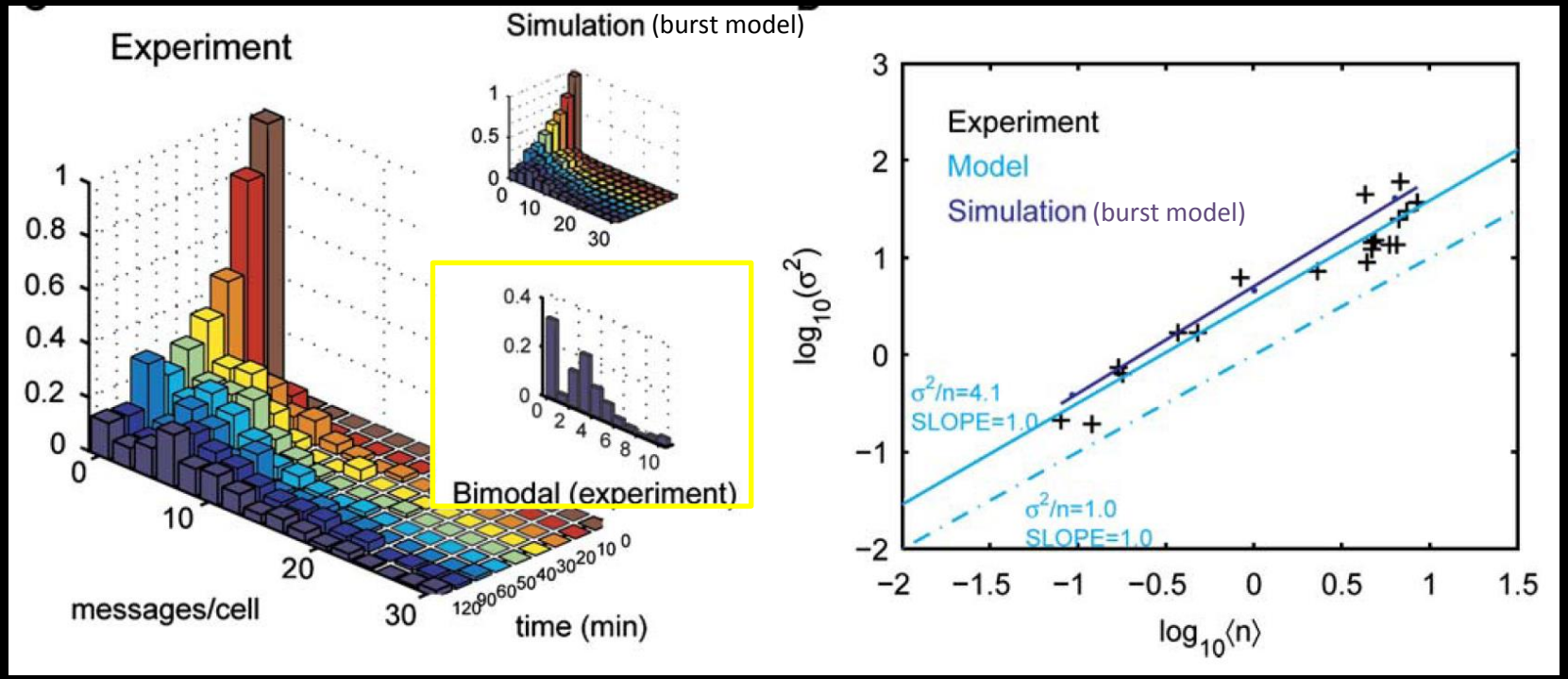
Probability to have no mRNA



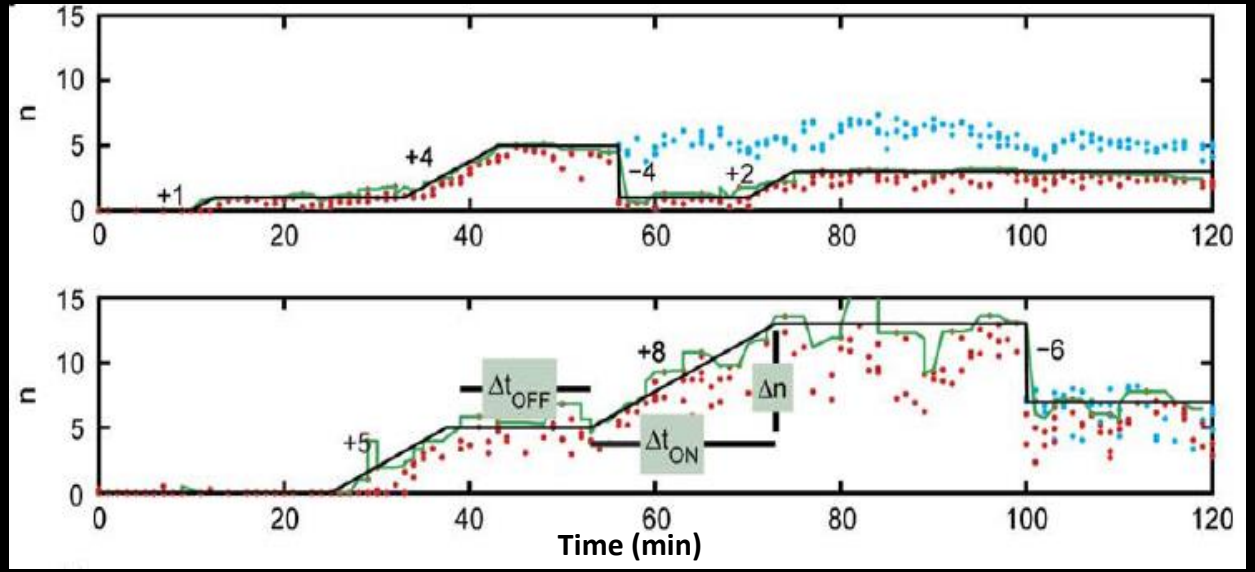
Not Poissonian!!

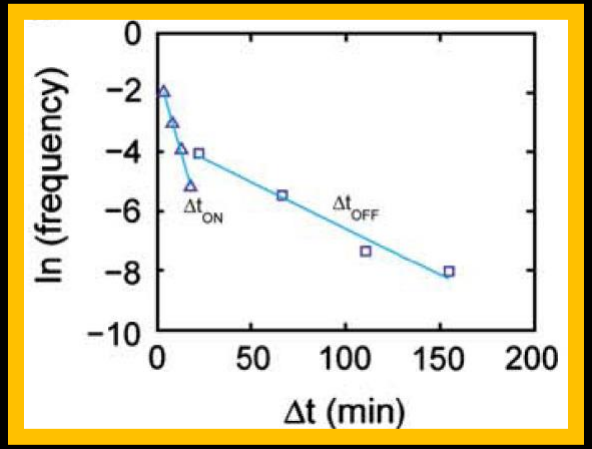
Testing the Poisson hypothesis 2

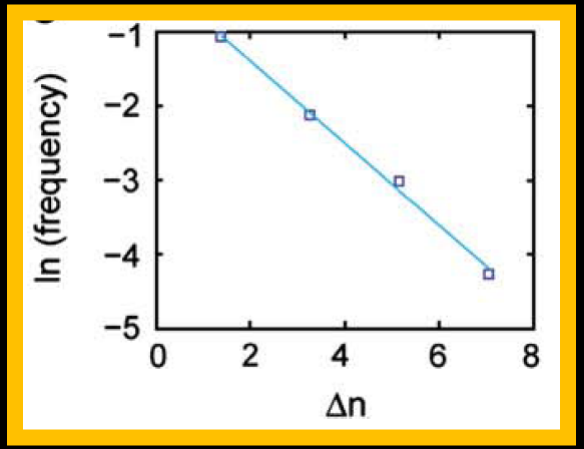
$$\sigma^2 = \langle n \rangle$$



mRNA synthesis in single cell

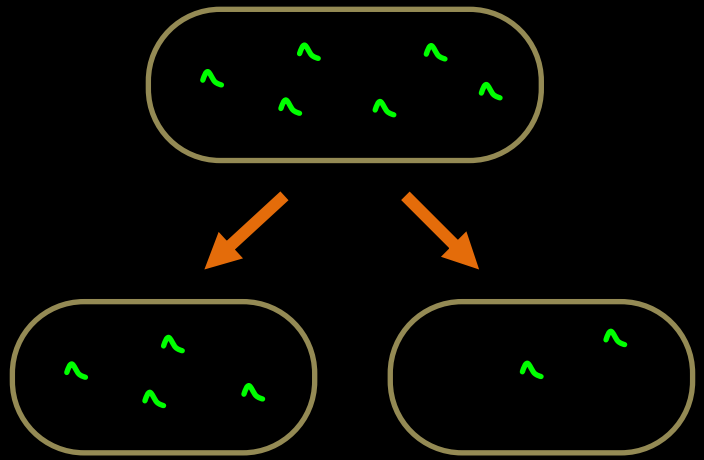






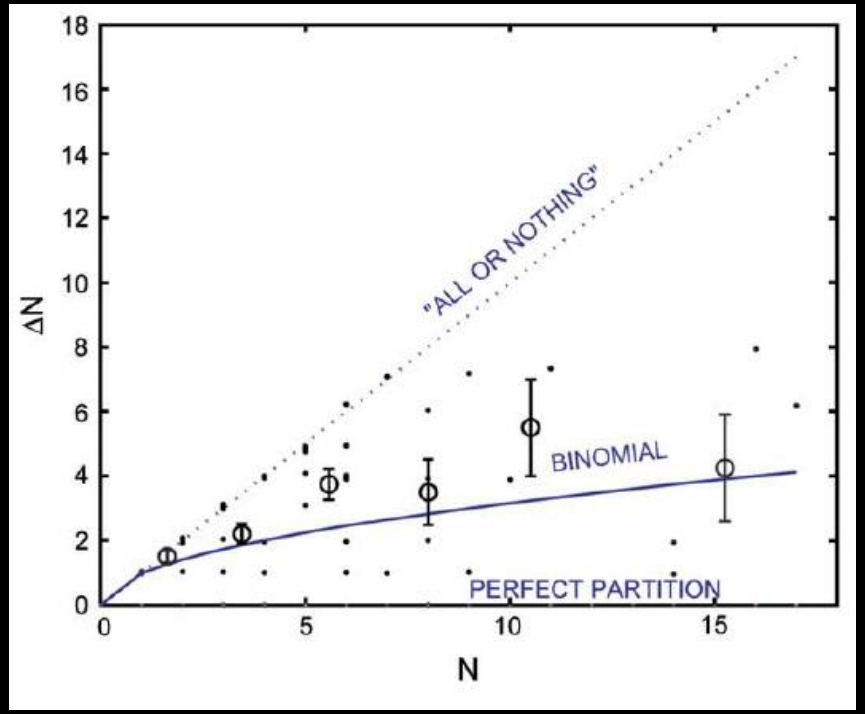
➔ **Burst of Transcription**

Cell division & partioning



$$\Delta n = \left| n_{left} - n_{right} \right|$$

$\Delta n \sim \sqrt{n}$, if random partion



Correlation between mRNA & protein

$$I_G = nNf_{GFP}$$

$$I_R = mf_{RFP}$$

$$\langle p \rangle = \left\langle \frac{m}{n} \right\rangle = N \times \left(\frac{f_{GFP}}{f_{RFP}} \right) \times \left(\frac{I_R}{I_G} \right) \approx 60 \sim 110$$