Reconstruction of transcriptional dynamics from gene reporter data

B. Finkenstadt[1,2], E. Heron[1,2], D. Rand[1,3], M. Komorowski[1,2,4]

***Short Abstract* — Transcriptional regulation is a fundamental process in cellular systems. In this study we address the generic question of reconstructing transcriptional dynamics from gene reporters data. This problem is important for understanding the complex regulatory mechanisms.**

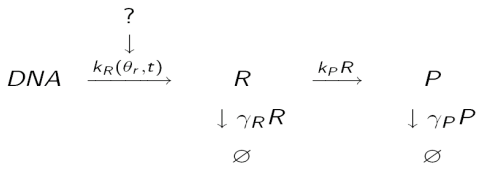
***Keywords* — Transcription reconstruction, reporter genes, Bayesian statistical inference for biochemical kinetics.**

I. Motivation

Recently, novel insight into gene expression dynamics has been obtained from single-cell experiments that draw upon the unique capabilities of fluorescent reporter proteins producing fluorescence roughly proportional to the actual number of fluorescent molecules in single-cells [1,2]. Experiments on single-cell level are necessary for investigating the stochastic nature of gene regulation in particular to infer transcriptional interactions between transcription factors and their target genes [2,3]. On the other hand, bulk-scale methods that measure average values over population also provide information about transcriptional dynamics nevertheless mask the behavior of individual cells and neglect implications provided by system fluctuations [4]. Different properties of population and single-cell experimental techniques impose conditions on statistical methods used for inference [5,6]. Here we introduce a general statistical methodology based on ordinary and stochastic differential equation models to address the generic problem of reconstructing transcription time course profiles from time series data on the reporter protein incorporating any other available information.

II. Modeling framework

Gene expression is a process involving dozens of biochemical reactions. Nevertheless it is commonly accepted to incorporate only four reactions (see Figure) into the mathematical model [7]. According to the size of observed molecular populations we use macroscopic rate equation or Fokker-Planck equation with corresponding Itô diffusion to draw inference about transcription profile $k_R(t,\theta_r)$, hidden mRNA process $R(t)$ and possibly other unknown parameters $(\gamma_R,\gamma_P,k,k_P)$ from series of protein measurements $(kP_{t_1},...,kP_{t_{1n}})$ and additional information



$\pi(\theta)$ (prior). According to the available data different assumptions about probabilistic model and parametric form of $k_R(t,\theta_r)$ are made. MCMC methods are used to sample from the posterior distribution of all unknowns.

Within the proposed framework we compare inference methods using biological data that are very heterogeneous in quality, experimental design and time resolution. We show what experimental setting is necessary to estimate each single parameter of the model (identifiability). Computational difficulties related to the sampling methods are also resolved.

III. Conclusion

The proposed statistical modeling methods can be used to reconstruct unobserved mRNA transcription profiles and estimate parameters of interests from reporter gene experiments. It contributes to the understanding of transcriptional dynamics of single genes and thus of whole regulatory mechanisms.

References

[1] J-Q Wu et al., (2005) Counting Cytokinesis Proteins Globally and Locally in Fission, Science 310 .
[2] D. Longo, J. Hasty, (2006) Dynamics of single-cell gene expression, Mol Syst Biol. 2.
[3] R-S. Wang, (2007) Inferring Transcriptional Interactions and Regulator Activities from experimental Data, Mol. and Cel. 24.
[4] Elowitz et al., (2000) A synthetic oscillatory network of transcriptional regulators, Nature 403.
[5] Golightly A. et al., (2005) Bayesian Inference for Stochastic Kinetic Models Using Diffusion Approximation, Biometrics 61.
[6] Carmen G., (2003) Parameter Estimation in Biochemical Pathways: A Comparison of Global Optimization Methods, Genome Res. 13.
[7] Thattai M. et al., (2001) Intrinsic noise in gene regulatory networks, PNAS 98.

[1]Systems Biology Centre, University of Warwick, Coventry, UK.
[2] Department of Statistics, University of Warwick, Coventry, UK.
[3]Department of Mathematics, University of Warwick, Coventry, UK.
[4]M.Komorowski@warwick.ac.uk