U Chicago-CNRS Collaboration Program, project TF_bursts: PhD fellowship

Stochastic transcriptional bursting in space and time

Summary of the project: In order to generate a functional organism, cell fate decisions must be taken at the right place and at the right time. Decades of genetic studies in *Drosophila* have dissected the gene regulatory networks responsible for the establishment of precise patterns of gene expression. In certain cases, network architecture reveals error correction mechanisms. However, recent live imaging studies show that transcription occurs in a random on/off manner called “bursts”. **The details of how transcription factors control bursting are unknown and are the focus of this proposal.** Our project is concerned with a particular class of TFs called pioneer factors, such as Zelda. To study the impact of TF pioneer factors on transcriptional bursting in single cells, we will employ quantitative live imaging of transcription (MS2/MCP) combined to an innovative signal processing approach, able to position each single polymerase initiation event for each nucleus *in vivo*. **With a view to understanding how order emerges from stochastic local events, both microscopic and coarse-grained mesoscopic models will be developed.** The micro-meso relationship will be rigorously studied in the framework of limit theorems for Markov processes.

Experimental approaches:

The experimentation will be performed in the Lagha lab (IGMM UMR CNRS 5535) in Montpellier and in the Reinitz lab in the University of Chicago. The phD student, whose major role will be the mathematical modeling, will also be familiarized with the experimental techniques and will contribute to data production and analysis in a proportion compatible with his major activity.To monitor the dynamics of transcriptional activation of development genes, we will employ the MS2/MCP labeling method [1,2], whereby an MS2 tag is recognized by an MCP-GFP detector protein. We envisage two types of imaging: on live embryos and on fixed embryos. Some quantitative aspects of live imaging of transcription (calibration, detection threshold etc) must be confirmed with a more sensitive method like single molecule FISH, a technique routinely employed in the Lagha lab. For image analysis of MS2 live movies, we will use readily available image analysis software developed by the Lagha and the Reinitz lab for 3D detection and segmentation of nuclei and their transcriptional activity.

Mathematical modelling approaches:

New mathematical models and data analysis methods will be developed in this project in the Radulescu Lab (LPHI UMR CNRS 5235) in Montpellier.

Using movies of Drosophila embryos as input, we developed a quantitative analysis methodology that allows, for the first time, to unequivocally determine each polymerase initiation event, for each single nucleus. This technique, based on numerical deconvolution of the MS2 signal, is an elegant alternative to other methods that predict only indirectly and statistically the sequence of events [1]. Furthermore, stochastic promoter models (continuous time, finite state Markov chains) are found in two steps: i) we use parametric multi-exponential estimations of the cumulative distribution function of the waiting times between successive transcription events; ii) we then use symbolic solutions to the inverse problem to find the parameters of the promoter models. This procedure permitted the complete characterization of stochastic transcription from a reporter construct of a gene involved in the dorso-ventral patterning of embryos, revealing several metastable promoter states and dynamical transitions between these states (Dejean et al, unpublished). This method will be improved in the thesis by extending it to non-stationary signals and by a rigorous treatment of interval censoring.

Models of stochastic transcription resulting from the quantitative analysis of the MS2 signal will be combined with models of control by transcription factors to generate spatially

extended segmentation models. This was done in the past using deterministic ODE versions of the gene circuit model [3] as well as deterministic models of transcription [4], but did not treat the stochastic aspects. Radulescu obtained numerous mathematical results on intermittent transcription (bursting) [5,9], but not for spatially distributed expression. In spatial transcription bursting models we will consider intermittent local production of mRNA in the positions of the nuclei, and regulator protein transport possibly leading to enhanced spatial concentration fluctuations (hubs) interacting with the local transcription dynamics. We will use two classes of models. “Microscopic” models will be based on the probabilistic description of the local transcription events and protein transport and will be studied by stochastic simulation (Gillespie algorithm, Brownian dynamics, hybrid algorithms [5]). The rates of stochastic transcription events will be related to the concentrations of TFs using biochemical and thermodynamic models developed in the Reinitz lab [4]. “Mesoscopic” models will be obtained by coarse-graining [6] from the “microscopic” models. In past work, Radulescu and Reinitz have already used coarse-graining to unravel error-correcting mechanisms that emerge from genetic interactions and control the precision of patterning [7]. This strategy will be further developed in this project, using different mathematical tools. We want to extend results from [9,8] on limit theorems to the spatial transcription dynamics. This will lead to simplified processes describing the stochastic dynamics of gene expression at mesoscale (at distances larger than the spacing between nuclei). We expect to obtain a panoply of models, including stochastic partial differential equations (SPDE), as well as hybrid, or non-local SPDEs taking into account singular source terms describing intermittent transcription. A first step towards the implementation of this program was recently made in the Radulescu lab by using the Liouville equation to justify PDE models of tissues [8]. Hybrid models will be obtained by adapting to the case of spatial heterogeneity the techniques developed in [9] for a well-stirred reactor. The analysis of this panoply of models will shed light into the complex but fundamental question of transcription dynamics in the context of precise cell fate decisions in an embryo.

[1] N. C. Lammers, V. Galstyan, A. Reimer, S. A. Medin, C. H. Wiggins, and H. G. Garcia. Multimodal transcriptional control of pattern formation in embryonic development. Proceedings of the National Academy of Sciences, 117(2):836–847, 2020.
[2] J Dufourt, A Trullo, J Hunter, C Fernandez, J Lazaro, M Dejean, L Morales, K N Schulz, C.Favard, M.M. Harrison, O. Radulescu, M. Lagha. Temporal Control of Transcription by Zelda in living Drosophila embryos, Nature Communications, 2018, 9 (1): 5194.
[3] Manu, S. Surkova, A. V. Spirov, V. Gursky, H. Janssens, A. Kim, O. Radulescu, C. E. Vanario-Alonso, D. H. Sharp, M. Samsonova, and J. Reinitz. Canalization of gene expression and domain shifts in the Drosophila blastoderm by dynamical attractors. PLoS Computational Biology, 5:e1000303, 2009. doi:10.1371/journal.pcbi.1000303 PMCID: PMC2646127.
[4] K. Barr, J. Reinitz, and O. Radulescu. An in silico analysis of robust but fragile gene regulation links enhancer lengthto robustness. PLoS computational biology, 15(11), 2019.
[5] G. C. Innocentini, A. Hodgkinson, and O. Radulescu. Time dependent stochastic mrna and protein synthesis in piecewise-deterministic models of gene networks. Frontiers in Physics, 6:46, 2018.
[6] O. Radulescu, A. N. Gorban, A. Zinovyev, and V. Noel. Reduction of dynamical biochemical reactions networks in computational biology. Frontiers in genetics, 3:131, 2012.
[7] S. Vakulenko, Manu, J. Reinitz, and O. Radulescu. Size regulation in the segmentation of drosophila: interacting interfaces between localized domains of gene expression ensure robust spatial patterning. Physical review letters, 103(16):168102, 2009.
[8] A. Hodgkinson, G. Uzé, O. Radulescu, and D. Trucu. Signal propagation in sensing and reciprocating cellular systems with spatial and structural heterogeneity. Bulletin of mathematical biology, 80(7):1900–1936, 2018.
[9] A.Crudu, A.Debussche, A.Muller, and O.Radulescu, Convergence of stochastic gene networks to hybrid piecewise deterministic processes, Annals of Applied Probability (2012) 22: 1822-1859.

Techniques: Confocal imaging; Data analysis (image analysis, statistics, programming in R, Matlab or Python); Mathematical modelling (continuous- and discrete-time Markov processes).

Candidate: Motivation, perseverance, rigor, creativity and curiosity. Prior knowledge in quantitative biology is a plus but not mandatory. Interest in mathematical and quantitative biology and expertise in programming are mandatory. Required diploma: A research master degree from a University, Engineering School or “Grande Ecole”, in a field relevant to the project (biology, mathematics, physics, engineering). Students with interdisciplinary profile are strongly encouraged to apply.

Environment: This project belongs to an interdisciplinary network training two phD students in collaboration with three labs: Mounia Lagha (IGMM UMR 5535), Ovidiu Radulescu (LPHI UMR 5235) and John Reinitz (U Chicago). The PhD student will be based in Montpellier and co-supervised by OR and ML but will travel regularly to Univ. of Chicago to work with the Reinitz lab, where a second PhD student will be working on a related project. The experiments will be primarily performed in the Lagha lab and the mathematical modelling in the Radulescu lab.

How to apply: Please send an email containing “PhD France-Chicago” in its subject, with your motivation letter, CV and reference names to mounia.lagha@igmm.cnrs.fr and ovidiu.radulescu@umontpellier.fr, before May 31, 2020.

Funding: 36 months salary + travel money, start date 01/10/2020, provided by CNRS.