Toward a dynamic multi-scale/level approach for Gene Regulatory Network inference

Arnaud Bonnaffoux*[1,2], Ulysse Herbach[3,4,5], Angélique Richard[5,6], élodie Vallin[5], Sandrine Giraud[5,6], Camilo La Rota[2], Pierre-Alexis Gros[2], and Olivier Gandrillon[7]

[1]Laboratoire de Biologie Moléculaire de la Cellule (LBMC) – CNRS : UMR5239, Institut national de la recherche agronomique (INRA) : UR5239, Université Claude Bernard - Lyon I (UCBL), École Normale Supérieure (ENS) - Lyon – ENS de Lyon 46 allée d'Italie 69364 LYON Cedex 07, France

[2]The CoSMo Company (CoSMo) – The CoSMo Company – 5 passage du Vercors, 69007 Lyon, France

[3]CNRS – Centre National de la Recherche Scientifique - CNRS, INSA - Institut National des Sciences Appliquées, L'Institut National de Recherche en Informatique et e n Automatique (INRIA), Université de Lyon – France

[4]Institut Camille Jordan (ICJ) – Ecole Centrale de Lyon, Université Jean Monnet - Saint-Etienne, INSA - Institut National des Sciences Appliquées, Centre National de la Recherche Scientifique - CNRS, Université de Lyon – France

[5]Laboratoire de Biologie Moléculaire de la Cellule (LBMC) – Centre National de la Recherche Scientifique - CNRS, Université de Lyon, École Normale Supérieure (ENS) - Lyon, Institut National de la Recherche Agronomique - INRA (FRANCE) – France

[6]UCBL – Université de Lyon – France

[7]Université de Lyon, Université Lyon 1, Centre de Génétique et de Physiologie Moléculaire et Cellulaire (CGPHIMC) – Centre national de la recerche scentifique – CNRS UMR5534, F-69622 Lyon, France, France

Résumé

Gene regulatory networks (GRN) play an important role in many biological processes, such as differentiation, and their identification has raised great expectations for understanding cell behavior. Many computational GRN inference approaches have been described, which are based on expression data but they face common issues such as data scarcity, high dimensionality or population blurring (Chai et al., 2014). We believe that recent high-throughput single cell expression data (see e.g. Pina et al., 2012 ; Shalek et al., 2014) acquired in time-series will allow to overcome these issues and give access to causality, instead of " simple " correlation, for gene interactions. Causality is very important for mechanistic model inference and biological relevance because it enables the emergence of cellular decision-making. Emergent properties of a mechanistic model of a GRN should then match with multi-scale (single cell/population) and multi-level (molecular/cellular) observations.

We will expose a GRN inference framework based on these assumptions. It follows three steps:

*Intervenant

- Node parametric inference. We have inferred the parameters from a stochastic mechanistic model of gene expression, the Random Telegraph model (Kim and Marioni, 2013), thank's to time-series single cell expression data from a population of chicken erythrocyte progenitor during their differentiation process (Gandrillon et al., 1999)
- Model reduction. This is mostly an ongoing work, and will make use of specific constraints applying to the network.
- The final step will consist in network inference constrained by dynamic multi-scale/level observations.

Mots-Clés: single cell, GRN, inference, stochastic